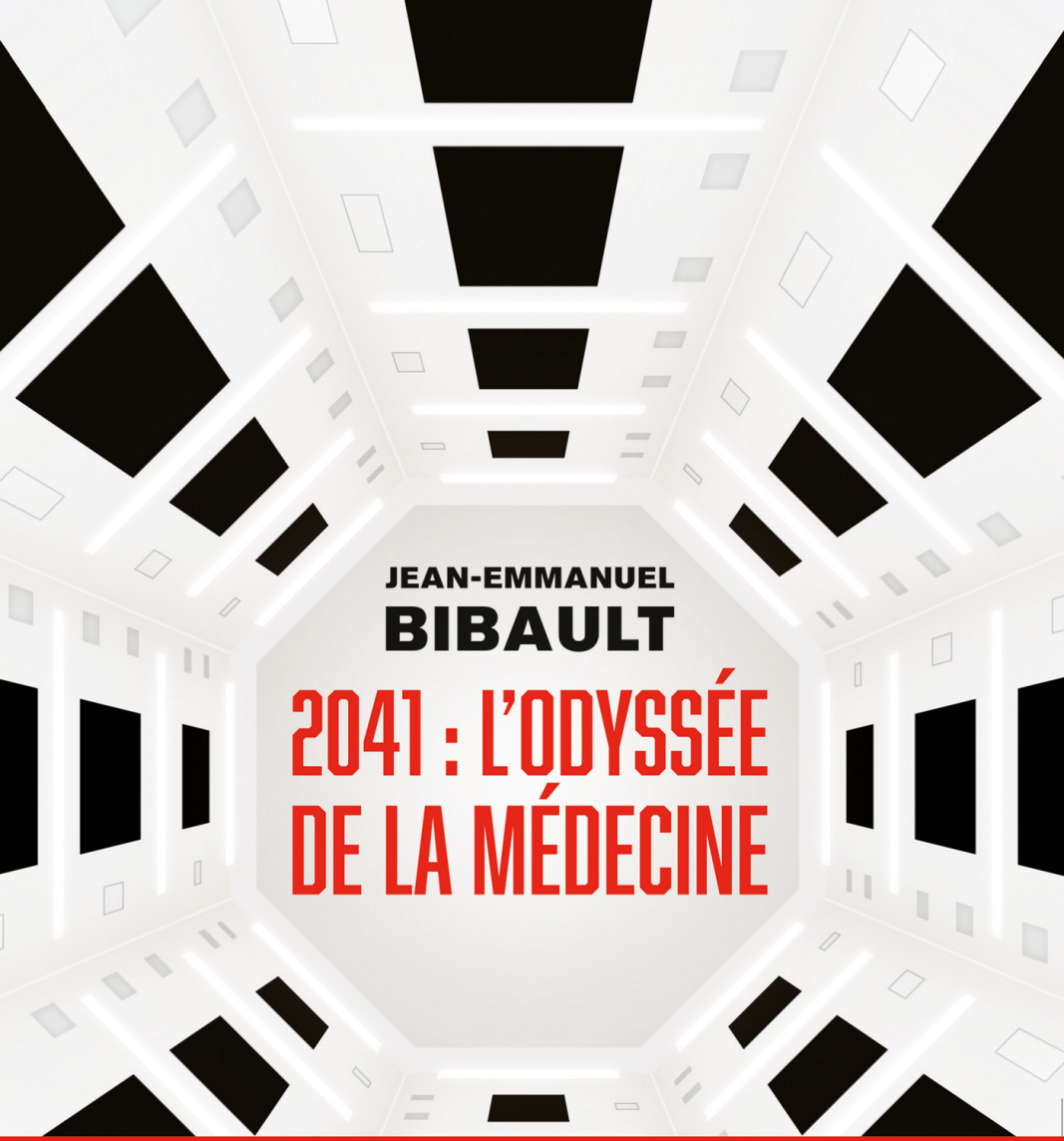
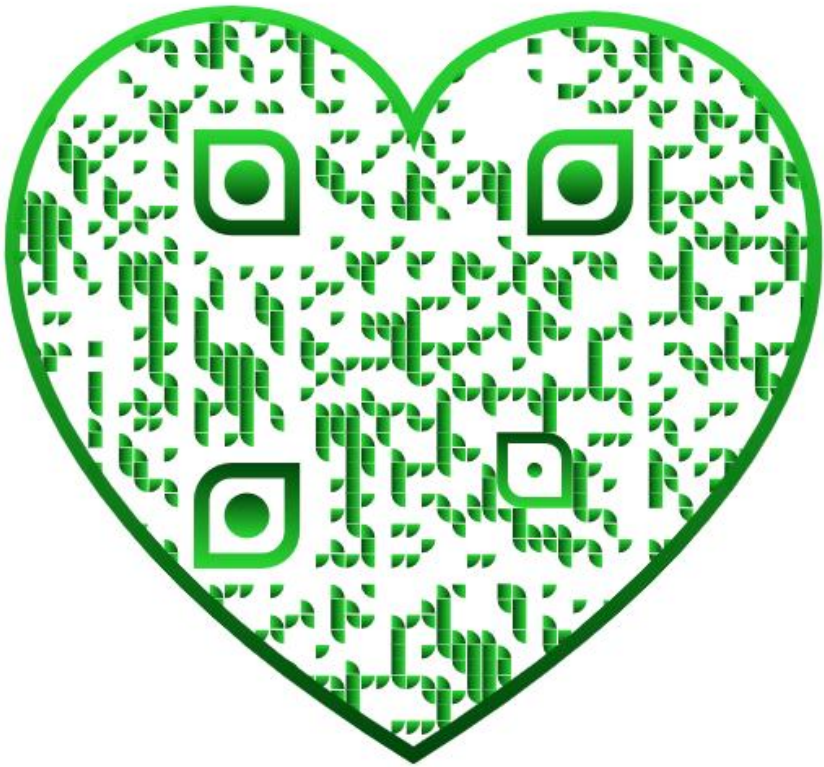


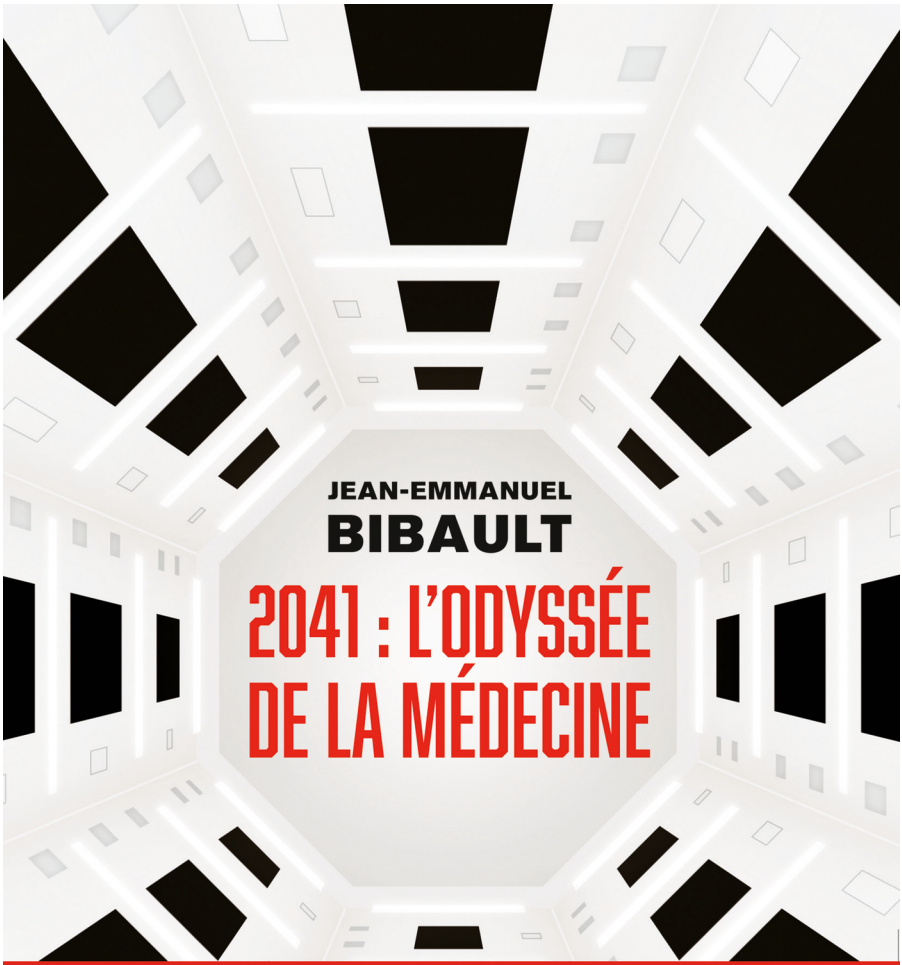
JEAN-EMMANUEL
BIBAULT

2041 : L'ODYSSÉE DE LA MÉDECINE

COMMENT L'INTELLIGENCE ARTIFICIELLE BOULEVERSE LA MÉDECINE

ÉQUATEURS





JEAN-EMMANUEL
BIBAULT

2041 : L'ODYSSÉE DE LA MÉDECINE

COMMENT L'INTELLIGENCE ARTIFICIELLE BOULEVERSE LA MÉDECINE

ÉQUATEURS

Jean-Emmanuel Bibault

2041 : L'ODYSSÉE DE LA MÉDECINE

Comment l'intelligence artificielle bouleverse la médecine

ÉQUATEURS

ISBN : 978-2-3828-4423-6.

Dépôt légal : janvier 2023.

© Éditions des Équateurs / Humensis, 2023.
170 *bis*, boulevard du Montparnasse, 75014 Paris.
editions-des-equateurs@orange.fr
www.editionsdesequateurs.fr

« La technophobie n'a rien à voir avec l'humanisme dont elle se réclame. « L'humanisme considère que l'homme est par essence un être technicien, un fabricant de machines – lui à qui la nature a si peu donné pour survivre sans artifice. »

Gilbert Simondon, *Du mode d'existence des objets techniques*, 1958.

Sommaire

Introduction

I **Contexte, histoire et concepts en intelligence artificielle**

1 La médecine à l'aube de l'intelligence artificielle
2 Entre mythes et philosophie
3 Qu'est-ce que l'intelligence artificielle ?
4 Brève histoire de l'intelligence artificielle et de son utilisation en médecine

II **Utiliser l'intelligence artificielle pour mieux soigner**

1 De la Silicon Valley et des robots
2 Dépister grâce à l'IA
3 Diagnostiquer et prévoir grâce à l'IA
4 Traiter grâce à l'IA
5 Prévenir grâce à l'IA
6 Le citoyen augmenté

III **Perspectives**

1 De la Pascaline au smartphone
2 Du creux de votre main à l'air qui vous entoure
3 La transformation du système de santé
4 Risques et dérives potentiels de l'intelligence artificielle appliquée à la santé
5 Secret médical et intelligence artificielle
6 Comment gérer et utiliser nos données ?
7 Faut-il réguler l'intelligence artificielle ?
8 Comment seront créées les intelligences artificielles de demain ?
9 Qui sera le leader en intelligence artificielle ?
10 2031 : une semaine, du dépistage au traitement d'un cancer du poumon
11 2041 : une IA remporte trois prix Nobel
12 Quel avenir pour la médecine ?

Annexes

Comment créer une intelligence artificielle ?
Comment fonctionne un réseau neuronal ?
Comment fonctionne un arbre décisionnel ?
Bases techniques du *Deep Learning*
À quoi reconnaît-on une bonne intelligence artificielle ?

Index des concepts

Remerciements

Introduction

Novembre 2020. Hillview Avenue, Palo Alto, Californie. Je me rends à pied à un rendez-vous chez Google Health, la division santé de Google, qui vient de s'installer à côté de notre laboratoire à Stanford. De nombreuses entreprises et les plus grands laboratoires d'intelligence artificielle siègent autour de l'université. C'est ici, au Xerox PARC, que Steve Jobs a assisté pour la première fois à une démonstration de la souris en 1979. L'importance de ce lieu pour l'histoire de l'informatique contraste avec la banalité de son architecture. Pas une plaque, pas un panneau ne le distingue des autres. De l'autre côté du 1070 Arastradero Road, où se trouve mon laboratoire de recherche, se dresse le siège de Tesla, le constructeur de voitures électriques dirigé par Elon Musk. Là aussi, la dissonance entre l'image renvoyée à travers le monde et le style du bâtiment est saisissante. Encore plus saisissante, la dichotomie totale entre l'émulsion intellectuelle qui doit se produire tout autour de moi et les rues désertes. Je suis le seul à marcher le long de ces grandes allées où passent quelques rares voitures. L'atmosphère des lieux et les quinze minutes de trajet me permettent de prendre un peu de recul sur mes projets et de repenser à ma trajectoire. Comment un Français, qui se destinait à la psychiatrie par rejet du déterminisme biologique, devient-il finalement cancérologue, chercheur en intelligence artificielle (IA) ?

Mon premier contact avec l'informatique remonte à la fin des années 1980, un peu avant mes 10 ans. Mes parents offrent un Amstrad CPC 6128 à mon frère et déjà nous programmons des jeux en BASIC en recopiant des lignes issues des manuels de l'époque (en toute honnêteté, ils ne fonctionnaient jamais). Quelques années plus tard, mes parents achètent un Apple LCII, principalement pour faire du traitement de texte, que nous utilisons surtout pour jouer, notamment au fameux jeu de course *The Duel* ! Au milieu des années 90, Internet arrive dans notre foyer. Notre immeuble parisien étant équipé du câble pour la télévision, notre connexion est bien plus rapide que celle de l'antique modem 56 K. Mes parents s'offrent donc un iMac Graphite Special Edition, du plus bel effet, dont je me souviens encore. En 1996, à 14 ans, je crée des sites Internet sur mon réalisateur préféré Stanley Kubrick et mon groupe de musique de l'époque. Je commence aussi à acheter des composants et à monter des PC sous Windows (véritable sacrilège dans la famille). Carte mère, RAM, processeur, disque dur et carte graphique n'ont bientôt plus aucun secret pour moi.

Après mon bac, j'entre à la faculté de médecine. Je considère à l'époque que tous les phénomènes régissant le corps humain – et donc susceptibles de le dérégler – ne peuvent s'expliquer simplement par la biologie, par des réactions chimiques ou physiques. Je conçois la médecine comme une

discipline éminemment artisanale. Ceux qui s'entêtent à vouloir tout quantifier, tout comprendre font forcément fausse route. Je me destine d'ailleurs à une discipline qui laisse encore une large place à l'inconnu et à l'interprétation, la psychiatrie. Mais, au cours de ma formation, je prends conscience qu'il est impossible de rejeter tous les progrès thérapeutiques accomplis grâce à la compréhension précise des mécanismes biologiques. Je suis aussi déçu de la faible efficacité des interventions thérapeutiques (médicamenteuses ou non) dans le domaine de la psychiatrie. En cinquième année, la découverte de la cancérologie et de la radiothérapie me fait l'effet d'une révélation. Ces disciplines allient technique, compréhension précise des traitements à administrer (grâce à la radiophysique et à la radiobiologie) et clinique. Par ailleurs la relation médecin-patient reste fondamentale. Je choisis l'oncologie radiothérapie pour mon internat à Lille. Parallèlement, je co-développe une application qui permet aux fumeurs de dépister les maladies en rapport avec le tabagisme chronique en répondant à une série de questions tous les mois. En comparant leurs réponses dans le temps, on peut détecter des signes associés au cancer, à la bronchite chronique ou à l'infarctus du myocarde.

En 2015, devenu chef de clinique en radiothérapie à Paris, j'entame une thèse de science. Sous l'impulsion de mes directeurs de thèse, les professeurs Burgun et Giraud, le thème de mes recherches s'oriente vers le *Machine Learning* appliqué à la cancérologie. Les différents articles que je publie, en rapport avec mes travaux, rencontrent un certain écho dans la communauté. En 2016, alors que j'assiste au congrès américain de radiothérapie à Boston, je m'aperçois qu'une partie du congrès est consacrée à ma thématique de recherche. Le responsable de la session, le professeur Lei Xing, de l'université de Stanford, cite même mes articles comme références durant sa présentation. Je le contacte afin de le rencontrer. Il me répond et me propose un rendez-vous au cours duquel je lui demande s'il serait possible pour moi de venir dans son laboratoire à Stanford afin de poursuivre mes recherches à l'issue de ma thèse. Il accepte et me propose un poste de chercheur post-doc. Je soutiens ma thèse de science en 2018 récompensée par un prix de l'Académie nationale de médecine. Comme prévu, je pars finalement en 2019 avec ma femme et ma fille à Stanford afin de poursuivre mes recherches sur l'intelligence artificielle appliquée à la cancérologie.

Devenu professeur en oncologie radiothérapie à l'université de Paris-Cité et praticien à l'Hôpital européen Georges-Pompidou, je poursuis mes recherches au Centre de recherche des Cordeliers dans le domaine de l'IA appliquée à la médecine.

Aujourd'hui, je suis convaincu que l'intelligence artificielle bouleversera profondément la façon d'exercer notre métier et améliorera le devenir des patients. Nous sommes en passe d'inventer les machines qui nous soigneront

mieux que nous sommes capables de nous soigner nous-mêmes. La machine se dressera bientôt contre la maladie, avec une implacable efficacité. Encore faut-il développer nos outils de la bonne façon.

I

Contexte, histoire et concepts en intelligence artificielle

1

La médecine à l'aube de l'intelligence artificielle

Pour comprendre pourquoi l'intelligence artificielle est en train de révolutionner la médecine, il faut avant tout comprendre ce qu'est la médecine au XXIe siècle. D'une façon globale, les principes qui régissent le dépistage, le diagnostic, le traitement et le suivi d'une maladie n'ont pas fondamentalement évolué depuis cinquante ans. Ainsi, le parcours d'un patient débute la plupart du temps par l'apparition d'un symptôme. La personne commence à ressentir une douleur, une gêne, une perte de poids, une fatigue. Le philosophe et médecin Georges Canguilhem écrivait très justement en 1943 : « La santé, c'est la vie dans le silence des organes. » Et c'est donc le plus souvent lorsque ce silence est rompu que la personne devra entrer dans une démarche diagnostique.

Mais avant que ces différents symptômes n'apparaissent, la maladie peut être déjà très évoluée, ce qui réduit les chances de guérison. À l'heure actuelle, malgré les importants progrès réalisés en imagerie ou en biologie médicale, le dépistage ou le diagnostic précoce des maladies reste complexe, voire impossible. Quand bien même une maladie pourrait être détectée à un stade très précoce (comme, par exemple, un cancer du pancréas), encore faudrait-il prouver que ce diagnostic précoce permettra d'améliorer la guérison ou la survie du patient. La capacité à détecter mais surtout à prédire l'évolution d'une maladie est essentielle dans la prise en charge. Or, à l'heure actuelle, la médecine n'est pas capable de fournir des modèles prédictifs qui soient efficaces et utilisables en routine clinique. Elle se contente, et c'est, il faut l'avouer, déjà extraordinaire, d'analyser un ensemble de symptômes présents afin de définir une maladie et donc le traitement correspondant. En ce sens, sur le principe même de la démarche médicale, l'intelligence artificielle permettrait non plus seulement de courir après la maladie mais d'en prévoir l'évolution et donc de l'anticiper pour mieux la combattre.

Prenons le parcours typique d'une patiente au XXIe siècle. Appelons-la Emma Langlois. 53 ans, sans antécédent médical particulier, qui fume un demi-paquet de cigarettes par jour depuis l'âge de 20 ans et commence à développer des difficultés à respirer et des douleurs thoraciques. Elle attendra probablement plusieurs mois avant de consulter son médecin traitant afin de lui parler de ses symptômes. La maladie aura déjà eu le temps d'évoluer. Après avoir trouvé le courage de consulter son médecin traitant, Mme Langlois devra lui exposer de la façon la plus fidèle possible la chronologie des événements qui ont mené à l'apparition de ses symptômes.

Le praticien développera alors des hypothèses diagnostiques en fonction de l'anamnèse (c'est-à-dire de son interrogatoire) et de l'examen clinique. Celui-ci, primordial, inclut l'inspection (observation du patient), la palpation (recherche d'anomalies comme des ganglions), la percussion (recherche de bruits anormaux comme par exemple une matité ou un tympanisme d'un organe) et l'auscultation (écoute des souffles cardio-vasculaires, bruits du cœur, de digestion, de la respiration avec un stéthoscope). Le médecin obtiendra alors une description générale de l'état clinique de la patiente. En fonction des anomalies constatées (et il peut aussi parfois n'y avoir aucune autre anomalie visible), le médecin débutera une démarche diagnostique paraclinique incluant des examens complémentaires. Dans le cas d'Emma Langlois, il est probable que l'on prescrive des explorations biologiques, fonctionnelles respiratoires ainsi qu'un scanner thoracique selon la gravité des symptômes. Imaginons maintenant que ce bilan révèle malheureusement une masse au niveau du poumon gauche. La patiente sera adressée à un médecin spécialiste afin de poursuivre les investigations.

L'étape suivante dans la démarche diagnostique consistera à réaliser une fibroscopie bronchique ou une ponction transthoracique afin d'obtenir une biopsie de la masse. Cette biopsie sera ensuite analysée sous microscope par un médecin anatomopathologiste afin d'en déterminer la nature. Si cette analyse révèle que la masse est un adénocarcinome pulmonaire, c'est-à-dire un cancer du poumon, il faudra faire un bilan d'extension : un scanner thoraco-abdomino-pelvien ou une tomographie par émission de positons au 18-FDG, au cours de laquelle on injecte un produit radioactif qui se fixe sur les cellules cancéreuses. Ces examens permettront de voir les organes touchés par la maladie. Ils nécessiteront malheureusement du temps, durant lequel la maladie continuera d'évoluer.

Le cancer de M^{me} Langlois se révélera finalement métastatique avec des lésions cérébrales millimétriques, des métastases hépatiques et osseuses. La patiente est alors adressée à un service spécialisé dans la prise en charge de ce type de maladie. Elle débute enfin son traitement plusieurs semaines voire plusieurs mois après l'apparition de ses premiers symptômes, même si les médecins et les équipes impliqués dans la réalisation du bilan initial font leur maximum pour aller au plus vite. Le traitement choisi sera celui qui aura montré sa plus grande efficacité d'après les études déjà publiées, une fois comparées les différentes approches. Ces études, très longues à mener (souvent une décennie ou plus), incluent en général plusieurs centaines, voire plusieurs milliers de patients, tirés au sort pour suivre un traitement standard et un autre. On compare ensuite les résultats de chaque traitement afin de retenir le meilleur pour l'ensemble des patients inclus.

C'est ici qu'entre en jeu l'intelligence artificielle. Comment peut-elle améliorer ce parcours et les traitements ?

Mener des études en médecine, comparer différents schémas de traitement ou différentes techniques peut se révéler difficile, voire impossible :

l'augmentation significative des paramètres cliniques et biologiques à intégrer dans un essai à l'heure de la médecine de précision complique la création d'essais dédiés. Il existe de plus une différence significative entre les populations de patients inclus dans les essais cliniques et les patients rencontrés dans la pratique clinique. Par exemple, les patients âgés, qui ne font quasiment l'objet d'aucun essai dédié, constituent une part importante de la population à traiter. De nouvelles approches sont nécessaires pour générer des études de bon niveau de preuve. Notamment, la création de systèmes d'intelligence artificielle entraînés à partir de bases de données massives détaillées, intégrant un grand nombre de paramètres. Ces systèmes doivent permettre d'accélérer le dépistage et le diagnostic des patients. Et aussi d'optimiser et de personnaliser leur traitement, afin d'augmenter les chances de guérison.

L'IA peut être utilisée à chacune des étapes du parcours du patient afin d'accélérer sa prise en charge et d'en optimiser les résultats. Que l'on soit cancérologue, cardiologue, dermatologue, ou médecin généraliste, l'IA est avant tout un formidable accélérateur de découvertes. Pourtant cette technique qui concentre tant d'espoirs nous a souvent déçus par le passé.

2

Entre mythes et philosophie

Les craintes habituellement suscitées par l'intelligence artificielle sont liées à la relation complexe qu'entretiennent les hommes avec les outils qu'ils créent.

Platon, dans *Le Politique*, débat des conséquences du remplacement des esclaves par les animaux pour accomplir des tâches physiques lourdes. La technique est ainsi soumise depuis longtemps à cette dialectique : il y a une contradiction indissociable entre les bienfaits que l'homme tire de son ingéniosité et les dangers qu'elle peut engendrer. Cela s'illustre parfaitement en médecine, où les progrès de la génomique permettent de mieux caractériser les maladies et de trouver de nouveaux traitements, mais risquent également de stigmatiser des individus. Il est ainsi possible de déterminer certaines caractéristiques d'une personne à partir de sa génétique ou de calculer son risque de développer une maladie. Avec de mauvaises intentions, on pourrait aussi utiliser le génome d'une personne pour lui imposer une destinée. Cet eugénisme est notamment décrit dans le film *Bienvenue à Gattaca*, réalisé en 1997 par Andrew Niccol, où les enfants conçus naturellement, avec de « mauvais génomes », sont contraints de devenir balayeurs, tandis que d'autres enfants, issus de l'ingénierie génétique, seront policiers ou astronautes.

On retrouve cette ambivalence vis-à-vis de la technologie également dans le mythe du Golem, qui date du XVI^e^ siècle : une « machine » imaginaire, un être artificiel et humanoïde, muet et dépourvu de libre arbitre, façonné à partir de la terre glaise par quatre sages figurant les quatre éléments (terre, air, eau et feu). Censé défendre la communauté des pogroms, il pouvait également échapper à son contrôle et se retourner contre elle.

La crainte de la technique est aussi celle du remplacement de l'homme par la machine. Cette crainte s'exprima par exemple à la fin du XVI^e^ siècle en Angleterre, lorsque la reine Élisabeth I^re^ rejeta William Lee et la machine à tricoter qu'il avait inventée, de crainte que celle-ci n'occasionne des pertes de savoir-faire et d'emplois parmi ses artisans. L'idée que l'outil remplace l'humain n'est donc pas nouvelle et l'étude réalisée en 2018 par une équipe d'Oxford, souvent citée, qui affirme que l'IA occasionnera la perte de 30 % des emplois actuels, n'a finalement rien d'étonnant et fait partie de l'éventail de fantasmes, plus ou moins justifiés, que l'IA suscite dans notre inconscient collectif : déshumanisation, surveillance et dictature robotique, guerre contre l'homme, pour n'en citer que quelques-uns. Ces craintes ont été largement nourries par le cinéma populaire depuis les années 1980.

Au XVII^e siècle, Descartes se fait l'avocat de la technique qui, explique-t-il dans le *Discours de la méthode* (1637), doit nous rendre « maîtres et possesseurs de la nature ». Il estime que le but ultime des sciences consiste à servir les hommes et que le service le plus utile est la médecine. Au siècle des Lumières, Jean-Jacques Rousseau avance que la civilisation, et donc indirectement la technique, a corrompu l'homme sauvage, pourtant bon par essence. Le progrès technologique a certes rendu la vie plus confortable mais il a aussi amolli les hommes physiquement et moralement. Dans *Émile ou De l'éducation* (1762), il écrit : « C'est notre industrie qui nous ôte la force et l'agilité. » De ce point de vue, l'utilisation de l'IA en médecine pourrait avoir l'effet inverse de celui recherché : en déléguant la démarche diagnostique et thérapeutique à des machines, nous perdrions le savoir et la capacité de nous soigner. À la même époque, Jean Le Rond d'Alembert, qui dirigea avec Diderot l'*Encyclopédie*, estime quant à lui que l'industrie fait le bonheur de l'humanité.

Emmanuel Kant, dans les *Fondements de la métaphysique des mœurs* (1785), considère que la technique doit être soumise à l'impératif moral. Puisqu'elle ne fixe pas le but des actions humaines, mais uniquement les moyens, la technique peut être utilisée à des fins bonnes mais aussi mauvaises. Il s'agit donc de toujours la subordonner à notre exigence morale. C'est une des raisons pour lesquelles les médecins doivent impérativement s'impliquer dans le développement de l'IA en médecine : personnellement engagés par le serment d'Hippocrate qu'ils ont prêté, ils sont les seuls à même de garantir le respect de ces valeurs déontologiques.

Henri Bergson a beaucoup écrit sur ce qui définit l'homme, et donc, par opposition, l'artificiel. Dans *L'Évolution créatrice* (1907), il écrit ainsi que « l'intelligence est la faculté de fabriquer des objets artificiels ». L'intelligence de l'homme est ce qui le définit en tant qu'espèce. Chez Bergson, l'intelligence correspond à notre faculté de modifier la matière, grâce à des outils, pour favoriser notre conservation. Nous n'aurions donc d'autres choix que d'en user à notre propre avantage, car elle serait enfouie en nous, comme un instinct impossible à réprimer. Il définit l'Homme non comme *sapiens* (qui sait), mais comme *faber* (qui fabrique). Au-delà de la connaissance, le geste fait l'humain.

La technique est source d'inquiétude et d'angoisse sociale. Cette façon de considérer la technique, l'invention et la mécanique a été renforcée par le romantisme du XIX^e siècle, notamment par le monstre de Mary Shelley dans *Frankenstein*. Au XX^e siècle, la science est mise au service du totalitarisme dans *Le Meilleur des mondes* (1932) d'Aldous Huxley et dans *1984* (1950) de George Orwell. Il n'y a guère que dans l'œuvre d'Isaac Asimov que l'on retrouve une description positive et bienveillante de l'IA et de la robotique, à travers notamment les trois règles qu'il a énoncées : 1. Un robot ne peut porter atteinte à un être humain ni, restant passif, laisser cet être humain exposé au danger ; 2. Un robot doit obéir aux ordres donnés par les êtres humains, sauf

si de tels ordres entrent en contradiction avec la première loi ; 3. Un robot doit protéger son existence dans la mesure où cette protection n'entre pas en contradiction avec la première ou la deuxième loi.

Dans *La Technique ou l'Enjeu du siècle* (1954) Jacques Ellul avance même que « la technique est devenue le mystère essentiel ». Elle rejoint le domaine du sacré, dans le sens où elle ne se laisse déterminer par rien d'autre, mais détermine tout : politique, économie, et donc aussi bien sûr médecine. Bien loin de la vision de Kant, elle est devenue un moyen et une fin en tant que telle, échappant à toute réflexion et à tout contrôle. C'est précisément ce que nous devons éviter dans le domaine médical. Il s'agit donc de persister, envers et contre tout, à « penser ce que nous faisons » comme l'affirmait Hannah Arendt (*Condition de l'homme moderne*, 1958). Cela suppose de ne pas laisser la question du progrès aux seuls experts du domaine, mais d'y intégrer un réel débat politique.

Les exemples dans le cinéma de science-fiction sont très nombreux et ambivalents. Dans *2001, l'Odyssée de l'espace* de Stanley Kubrick, sorti en 1968, une IA nommée HAL décide de tuer l'équipage du vaisseau spatial qu'elle pilote afin d'accomplir sa mission. On raconte d'ailleurs qu'Asimov, cité plus haut, sortit de la salle lorsque HAL tua le premier membre de l'équipage, enfreignant ainsi sa première loi de la robotique. Dans *Westworld* (Michael Crichton, 1973), des robots massacrent les hôtes d'un parc d'attractions. L'androïde Ash tente de tuer Ripley, l'héroïne d'*Alien* (Ridley Scott, 1979), mais dans l'opus suivant l'androïde Bishop se sacrifie pour aider et sauver les héros. Les *replicants* de *Blade Runner* (Ridley Scott, 1982), robots aux capacités surhumaines, sont interdits sur Terre et chassés pour être éliminés. Dans *Terminator* (James Cameron, 1984), un cyborg est renvoyé dans le passé pour tuer la mère du futur chef de la résistance humaine. Dans *Matrix* (Lana & Lilly Wachowsky, 1999), IA et robots réduisent l'humanité tout entière à l'état de simples piles électriques.

Les illustrations de la relation ambiguë que nous entretenons avec nos machines ne manquent pas. Si nous sommes à l'aube d'une nouvelle ère, où nous utilisons nos machines pour nous soigner, nous craignons tout autant qu'elles finissent par nous nuire. L'IA peut affranchir l'homme de sa condition, mais aussi le corrompre, voire l'avilir et le soumettre. Afin de mener à notre tour une réflexion pertinente sur l'IA en médecine, il nous est nécessaire d'en appréhender et d'en comprendre la définition, l'histoire, les techniques et les applications.

3

Qu'est-ce que l'intelligence artificielle ?

Il n'est pas simple de donner une définition claire de l'intelligence artificielle car le concept même d'intelligence a beaucoup varié en fonction des époques et des progrès en neuroscience. Deux grandes notions cependant nous ont été fournies par les précurseurs du domaine, à savoir Alan Turing et Marvin Lee Minsky. Alan Turing était un mathématicien et cryptologue britannique. Durant la Seconde Guerre mondiale, il travaille notamment en cryptanalyse sur la machine *Enigma*, un dispositif utilisé par les armées allemandes pour communiquer de façon secrète. Après la guerre, il œuvre à la conception d'un des tout premiers ordinateurs et contribue à la naissance du concept d'intelligence artificielle. En 1950, il présente le test de Turing dans son article fondamental *Computing machinery and Intelligence*. Le test propose de faire converser un humain avec un autre humain et un ordinateur. Si la personne n'est pas capable de distinguer l'ordinateur d'un humain, le test est un succès et on peut dire que la machine est douée d'intelligence. Ce test a été réalisé avec un succès relatif dans différents contextes, notamment dans le domaine médical, comme nous le verrons plus tard.

Mais l'expérience ne suffit pas à caractériser une intelligence artificielle. Pour une définition plus généraliste, il faut se tourner du côté de Marvin Lee Minsky, cofondateur du groupe d'intelligence artificielle du Massachusetts Institute of Technology (MIT). Pour lui, l'intelligence artificielle correspond à « la construction de programmes informatiques qui s'adonnent à des tâches [...] demandant des processus mentaux de haut niveau tels que l'apprentissage perceptuel, l'organisation de la mémoire et le raisonnement critique ». L'IA correspond donc à la capacité de percevoir un phénomène, raisonner pour le comprendre et s'adapter à de nouvelles situations.

Actuellement, par abus de langage, on appelle IA toutes les techniques liées à l'apprentissage machine, le *Machine Learning*. Au sein de ce champ existe également le *Deep Learning* qui a permis les progrès les plus récents. En réalité, l'IA ne se réduit pas au *Machine Learning*.

On distingue l'IA dite symbolique et l'IA issue de l'apprentissage machine. L'IA symbolique correspond à un algorithme construit à partir des connaissances d'experts. Dans le domaine médical, l'IA symbolique correspond par exemple à la création d'un algorithme de règles successives qui vont s'appliquer à un patient. Ces règles sont définies par un humain et correspondent à un ensemble de questions qui permettent de classer correctement une information. On répond à chaque question dans un ordre

prédéfini et l'on obtient une classification. À l'opposé, l'IA résultant de l'apprentissage machine a été entraînée sur de larges bases de données, sans règles préalables, et les définit elle-même. Ce type d'IA « apprend par elle-même » à partir des données, sans intervention humaine.

On distingue également l'IA *faible* de l'IA *forte*. Une IA *faible* est une intelligence artificielle qui se concentre exclusivement sur une tâche bien précise. Une IA *faible* entraînée pour reconnaître un chien ou un chat sur une image sera totalement incapable de reconnaître une embolie pulmonaire sur un scanner thoracique. À l'opposé, une IA *forte* devrait être capable de réaliser un grand nombre de tâches différentes. On parle alors également d'IA *générale*. La recherche de ce type d'IA s'assimile encore à une quête du Graal, à cause des contraintes techniques liées à l'apprentissage de réseaux. En effet, si une IA entraînée à une tâche est réentraînée pour en accomplir une autre, elle perd alors (ou oublie) sa capacité à exécuter la première (ce que l'on appelle le *catastrophic forgetting*). Cet oubli est lié schématiquement à la « reprogrammation » du réseau neuronal[1] nécessaire à l'exécution de la nouvelle tâche.

Si l'on s'intéresse maintenant aux différentes techniques d'apprentissage machine, il est nécessaire de connaître quelques grandes catégories. Parmi ces techniques, la plus fréquemment utilisée est celle de l'*apprentissage supervisé* : des données labellisées sont fournies à un algorithme et ce dernier apprendra tout seul à les reconnaître en fonction de leurs caractéristiques et de leurs motifs intrinsèques. L'intelligence artificielle apprend alors de chaque exemple, en ajustant ses paramètres, de façon à réduire l'écart, c'est-à-dire l'erreur, qui existe entre le résultat qu'elle donne et les valeurs réelles. On dit alors que l'IA converge vers une erreur minimale. Elle calibre ainsi les paramètres qui permettent d'obtenir le bon résultat au fur et à mesure.

À l'opposé, au cours de l'*apprentissage non supervisé* les données fournies ne sont pas labellisées. Par exemple, des radiographies de thorax sans le diagnostic qui les accompagne. L'IA cherche alors elle-même les motifs intrinsèques aux données, sans savoir ce qu'elle « cherche ». Ce type d'approche pourrait avoir un intérêt particulier en médecine, afin par exemple de comprendre de nouveaux mécanismes physiopathologiques, ou de prédire des structures moléculaires inconnues.

La particularité de l'IA *connexionniste* (reposant sur des réseaux de neurones artificiels) a longtemps été que, de par son auto-apprentissage, nous n'étions pas capables de comprendre les raisons pour lesquelles elle réalisait une prédiction. Depuis une dizaine d'années, de nouvelles techniques, dites d'interprétabilité, permettent de surmonter ces difficultés et de comprendre ou de visualiser les caractéristiques et les motifs qui ont permis à l'IA de répondre à une question. Cette interprétabilité revêt une importance particulière dans le domaine médical où la confiance en l'intelligence artificielle est la condition première de sa généralisation.

4

Brève histoire de l'intelligence artificielle et de son utilisation en médecine

Pour comprendre les répercussions de l'automatisation par intelligence artificielle en médecine, il faut en reprendre l'histoire. Il est très difficile de distinguer l'histoire de l'intelligence artificielle de celle de la robotique, tout comme il est parfois impossible de faire la part des choses entre inventions réelles, mythes ou escroqueries.

Le concept de machines automatiques émerge dès l'Antiquité avec les colosses de bronze et les servantes d'or décrites dans la mythologie. Le premier véritable automate de l'Histoire aurait été créé il y a plus de deux mille ans par Archytas, philosophe, mathématicien et astronome grec. Il aurait inventé une colombe, ou un pigeon, en bois, capable de voler grâce à de l'air comprimé ou de la vapeur. Un peu plus tard, Héron d'Alexandrie (Ier siècle après J.-C., Égypte) inventera plusieurs machines à vapeur permettant de distribuer automatiquement de l'eau ou d'ouvrir des portes. Il aurait également mis au point une machine parlante pour rendre les oracles, qui était en fait une escroquerie.

Dans le monde entier, d'autres automates et machines sont inventés : l'horloge publique à eau de Su Song (1020-1101, Chine) ou les automates d'Al Jazari (1136-1206, Turquie). En Occident, au XIIIe siècle, Roger Bacon crée un « homme de fer », Albertus Magnus met au point une statue parlante, capable de réaliser des prophéties. Ces automates sont très mal perçus : la tête parlante d'Albertus Magnus aurait ainsi été détruite par Thomas d'Aquin, effrayé par une telle abomination ! À la Renaissance, Léonard de Vinci invente lui aussi des machines en s'inspirant du vivant.

Au XVIIe siècle, les automates se développent au Japon avec les poupées de l'ère Edo qui peuvent servir le thé ou tirer à l'arc. Il est intéressant de noter que, contrairement à la culture occidentale, les automates n'ont jamais été craints au Japon, ce qui explique leur développement et leur acceptation dans cette culture jusqu'à aujourd'hui.

Plus tard au XVIIIe siècle, Jacques de Vaucanson invente un canard digérant. Ce premier robot, conçu vers 1734, très réaliste et complexe, est capable de simuler la digestion et la défécation.

Le projet initial de son créateur était de construire des machines capables d'imiter le vivant et donc indirectement d'en comprendre le mécanisme en s'appuyant sur la théorie de l'animal-machine de Descartes. S'il existait des automates « qui eussent des organes et la figure d'un singe, ou de quelque

animal sans raison, nous n'aurions aucun moyen pour reconnaître qu'[ils] ne seraient pas en tout de même nature que les animaux ». Il s'agit donc de créer un être artificiel qui ne serait pas distinguable d'un être vivant. Ce principe, transposé à l'intelligence, induit la possibilité de créer de toutes pièces un être intelligent, à condition d'expliquer et de reproduire chaque système neurologique à l'origine d'une perception ou d'une décision. En ce sens, cette notion est très proche de l'intelligence artificielle *connexionniste* : reproduisons les réseaux neuronaux de nos cerveaux et nous aurons reproduit l'intelligence. La description que Vaucanson fait de son canard est particulièrement impressionnante pour l'époque : « Il boit, barbote dans l'eau, croasse comme le canard naturel [...], il allonge son cou pour aller prendre du grain dans la main, il l'avale, le digère, et le rend par les voies ordinaires tout digéré. » Le robot est montré pour la première fois en 1739 à Paris en même temps que deux autres automates, le flûteur et le tambourinaire, devant mille cinq cents spectateurs. Le succès est tel que Voltaire salue Vaucanson et le qualifie même de nouveau Prométhée. Par la suite, le canard changera de propriétaire à de nombreuses reprises et subira des dégradations et des réparations jusqu'à sa disparition, plus d'un siècle plus tard en 1879. À la même époque, l'horloger Pierre Jaquet-Droz met au point trois automates : la musicienne, le dessinateur et l'écrivain. Ils soufflent, respirent, tournent la tête et les yeux, comme de vrais humains.

En 1844, Joseph Faber dévoile à New York, puis à Philadelphie, le fruit de vingt-cinq ans de travail : Euphonia, une tête capable de parler français, anglais et allemand. L'échec public est tel que l'inventeur détruit sa machine dans un accès de désespoir. Il la reconstruira pourtant l'année suivante et rencontrera enfin le succès, grâce à l'intervention d'un showman américain, P.T. Barnum, qui en fera la promotion. La machine fut examinée et son fonctionnement authentifié par un scientifique américain, Joseph Henry, afin de vérifier qu'il ne s'agissait pas d'une fraude, comme c'était souvent le cas à cette époque. L'expert rapporta que « le plan de la machine [était] le même que celui des organes humains de la parole, les différentes parties étant travaillées par des cordes et des leviers au lieu de tendons et muscles ». Euphonia est, d'une certaine façon, l'ancêtre de Siri et d'Alexa.

Entré dans les annales en 1770, le Turc mécanique construit par Johann Wolfgang von Kempelen est particulièrement intéressant dans l'histoire de l'intelligence artificielle. Cet automate donnait l'illusion de jouer aux échecs contre un adversaire humain. Habillé d'une cape et coiffé d'un turban, il était assis derrière un meuble censé contenir des mécanismes et des engrenages. En réalité, un joueur humain bien réel, dissimulé dans un compartiment secret, manipulait le mannequin, tel un marionnettiste.

L'anecdote nous rappelle l'importance de conserver un esprit critique vis-à-vis des éventuelles prouesses plus ou moins irréalistes que l'on rencontre parfois dans le domaine de l'intelligence artificielle.

Il faudra attendre plusieurs siècles avant que Leonardo Torres, ingénieur

et mathématicien, invente le premier automate joueur d'échecs, en 1914. Cet automate était capable de jouer la finale roi et tour contre roi seul face à un adversaire humain, avec un bras mécanique qui déplaçait les deux pièces. Son algorithme parvenait toujours à mettre le roi du joueur humain en échec. Ce fut à cet égard le premier automate capable de prendre une décision automatiquement et le premier jouet électronique.

En 1920, le terme « robot » apparaît pour la première fois, dans la pièce de théâtre *Rossumovi univerzální roboti*, de l'auteur tchécoslovaque Karel Čapek. Le mot, inventé par son frère Josef, provient du mot *robota*, qui signifie corvée en tchèque.

En 1943, aux États-Unis, le neurophysiologiste Warren McCulloch et le psychologue Walter Pitts conçoivent une modélisation mathématique d'un réseau neuronal. Cette vision *connexionniste* de l'intelligence artificielle dominera toute la première période cybernétique du développement de l'IA. McCulloch et Pitts imaginent un neurone qui prend des variables en entrée (par exemple les caractéristiques d'un patient) et y applique un poids pour calculer une somme. Si cette somme dépasse un certain seuil, le neurone sera activé[2]. Le poids représente schématiquement l'importance d'une caractéristique sur la prédiction. Si le poids est faible, la caractéristique n'est pas importante et ne participe pas à la prédiction et inversement. À l'époque, les inventeurs des réseaux neuronaux ne parlaient d'ailleurs pas d'intelligence artificielle. Ce modèle du neurone, inspiré du vivant, deviendra pourtant la pièce centrale du développement de l'IA. Dans ce concept de machine intelligente, l'information est un signal, une forme pure indépendante de sa stricte compréhension. La machine ne fait qu'apprendre et réagir à l'information, sans la comprendre. Une des applications initiales de cette théorie fut utilisée en 1948 pour un système de guidage de missiles antiaériens par Norbert Wiener. Ce système reposait sur la mise à jour en continu de la trajectoire du missile en comparant le trajet réellement effectué à celui initialement prévu, ce qui permettait de corriger et d'orienter en temps réel la trajectoire. La machine s'adaptait à l'environnement et n'avait pas besoin d'en avoir une représentation ou une compréhension propre.

À l'opposé du domaine de l'armement, la cybernétique est également appliquée à la compréhension de la physiologie par W. Ross Ashby, un psychiatre et ingénieur anglais, qui a inventé l'Homéostat en 1946.

La machine est composée de quatre caissons interconnectés remplis d'eau. Au sein de ces caissons, des palettes mobiles occupent une position fixe d'équilibre générée par des potentiels électriques qui s'annulent mutuellement. Une intervention extérieure peut perturber ce système temporairement. Grâce à l'influence des éléments du système les uns sur les autres, les palettes retrouvent spontanément leur état d'équilibre malgré la perturbation. Cette machine, capable de s'adapter à son environnement de façon autonome, illustre parfaitement le principe fondamental de la cybernétique.

À la fin des années 1950, les réseaux de neurones connaissent un certain essor. De façon simplifiée, un réseau neuronal est un ensemble de couches de neurones. Chacun de ses neurones prend les valeurs d'entrée qui lui sont fournies, y applique une règle qui lui est propre et active ou non le neurone suivant, ce qui permet la transmission de l'information. Chaque neurone est une petite unité indépendante, comme dans un vrai cerveau, qui effectue certaines opérations et passe l'information aux neurones auquel il est connecté. Entre les couches d'entrée et de sortie, d'autres couches sont capables d'apprentissage.

Le terme même d'intelligence artificielle est inventé en 1955, lors d'un séminaire organisé par John McCarthy et Claude Shannon à Dartmouth. Cette dénomination et ce concept s'opposent à l'époque au *connexionnisme*. Il ne s'agit plus seulement de laisser une machine s'ajuster automatiquement à son environnement mais plutôt de lui donner la possibilité de le faire en manipulant et en appliquant des règles préétablies. C'est la naissance de l'IA *symbolique*. Mais afin que cette IA *symbolique* puisse exister, encore faut-il se montrer capable de créer des règles et des représentations de haut niveau. Cette théorie de l'esprit sera poussée par Minsky bien plus tard en 1986.

En 1957, Frank Rosenblatt, psychologue et informaticien de l'université de Cornell (États-Unis) crée le Perceptron, une machine capable de « voir » des images et de les classer en fonction de ce qu'elles contiennent. Le Perceptron est constitué de plusieurs couches de neurones : les neurones d'entrée correspondent à la rétine et les neurones de sortie classent les images reconnues par le système. Cet appareil est en quelque sorte l'ancêtre des réseaux neuronaux convolutifs qui nous permettent aujourd'hui de diagnostiquer automatiquement des maladies sur des scanners, ou encore de prédire la prévalence du cancer à partir d'images satellite, comme je l'expliquerai dans les chapitres suivants. Frank Rosenblatt a d'abord conçu un premier prototype logiciel sur IBM 704, puis construit une version mécanique du Perceptron, le Mark 1, avec 400 cellules photoélectriques. Les poids de chaque neurone étaient enregistrés dans des potentiomètres et les variations de poids (lors de l'apprentissage) produites par des moteurs électriques.

Cette approche domine le champ de l'intelligence artificielle de la fin des années 1950 au début des années 1990. Les Américains Herbert Simon et Allen Newell écrivent le premier programme capable de simuler une prise de décision, le *Logic Theorist*, en 1956. Les deux ingénieurs annoncent même à l'époque que « d'ici Noël [ils] auront inventé une machine pensante ». Cette stratégie de recherche se cristallise autour d'un groupe comprenant le Massachusetts Institute of Technology (MIT) avec Minsky et Papert, l'université Carnegie Mellon avec Simone et Newell et l'université de Stanford avec McCarthy. Ils parviendront à capter 75 % des financements de recherche de la Defense Advanced Research Projects Agency (DARPA), l'Agence pour les projets de recherche avancée de défense, et de l'Armée de l'air américaine, grâce aux applications militaires potentielles d'une

intelligence artificielle efficace.

Dans le même temps, l'intelligence artificielle dite *connexionniste*, perçue comme une boîte noire dont on ne comprend pas bien le fonctionnement, est sévèrement critiquée. Seymour Papert, chercheur au MIT, démontre par ailleurs que les réseaux de neurones sont incapables de traiter des phénomènes non linéaires – qui ne se produisent pas de façon directement proportionnelle à l'entrée – et sont par conséquent inutilisables dans la vie de tous les jours. En 1971, Frank Rosenblatt, l'inventeur du Perceptron, décède prématurément, entraînant avec lui les réseaux neuronaux qui disparaissent du champ de la recherche en IA. Dans les années 1970, l'IA entre dans son premier « hiver ». Les résultats de l'IA *connexionniste* ou *symbolique* ne sont pas à la hauteur des annonces faites à la fin des années 1960. Progressivement, tous les soutiens financiers des différents projets se retirent. En 1972, le philosophe américain Hubert Dreyfus explique dans son ouvrage *What computers can't do* que les règles ne suffisent pas à rendre des machines intelligentes.

Il faudra attendre quasiment une décennie avant que l'IA ne renaisse de ses cendres à travers la création de « systèmes experts ». Ces machines possèdent alors un très grand répertoire de connaissances traduites et incorporées depuis le savoir d'experts. À partir d'algorithmes simples composés de règles telles que « si, alors », l'utilisateur est invité à répondre à un certain nombre de questions afin d'obtenir une prédiction, par exemple une classification. Ainsi, en fournissant un certain nombre de symptômes, on peut obtenir une prédiction de la maladie qui en est la cause. Dans le domaine médical, cette approche semble particulièrement séduisante puisque l'on peut par exemple, à partir d'une liste de symptômes, déduire les maladies les plus probables. Les financements reviennent. Les médias annoncent à nouveau l'arrivée imminente de machines intelligentes. Mais, afin de tenir compte de l'ensemble des possibilités du monde réel, les chercheurs en IA sont contraints de concevoir des algorithmes extrêmement complexes, entraînant une inflation de règles incontrôlable et irréaliste.

Cette seconde vague des années 1980 cherche à créer une représentation du réel la plus exhaustive possible par la création d'ontologies générales, sortes de traduction et de classification totale du monde, pour le rendre objectivable dans un système fini informatique. Cette approche de l'intelligence artificielle est quasi binaire : un élément est vu comme vrai ou faux. La complexité des domaines abordés par la recherche en intelligence artificielle rend ce déterminisme illusoire et c'est à cette époque qu'émergent des notions plus probabilistes, moins tranchées.

Au début des années 1990, on comprend qu'il est illusoire de vouloir créer un nombre infini de règles explicites afin de représenter le monde. Le concept d'IA *symbolique* disparaît progressivement.

Fort heureusement, plusieurs groupes ont poursuivi leurs recherches dans le domaine des réseaux neuronaux. En juin 1979, à La Jolla en Californie, une réunion est organisée par le Canadien Geoff Hinton et l'Américain James

Anderson afin de discuter des possibilités d'autres approches de la recherche en IA. Sept ans plus tard, ce groupe publiera ses travaux, où réapparaissent les notions de neurones et de réseaux, sous le titre *Parallel Distributed Processing* (PDP). Les chercheurs travaillent sur de nouveaux types de réseaux neuronaux multicouches. Ils mettent également au point une méthode de rétropropagation (*backpropagation*) qui permet d'entraîner un réseau de neurones profond. On peut ainsi faire remonter dans les couches neuronales l'erreur constatée entre la prédiction réalisée et la réalité depuis la sortie jusqu'à l'entrée du réseau. Pour certains, cette *backpropagation* a en fait été inventée en 1970 par le mathématicien finlandais Seppo Linnainmaa. En réalité, il est probable que plusieurs chercheurs aient inventé la même chose, chacun de leur côté !

En 1989, l'informaticien français Yann Le Cun entraîne un réseau neuronal à reconnaître des chiffres manuscrits. Cette innovation majeure, réalisée aux Bell Labs, laboratoire fondé par Graham Bell, l'inventeur du téléphone, pose les bases du *Deep Learning*. Petit à petit, on diversifie les fonctions d'activation des neurones, on augmente le nombre de couches des réseaux et on complexifie leurs architectures.

De l'autre côté de l'Atlantique, une équipe de l'université de Munich met également au point l'un des tout premiers réseaux neuronaux profonds au début des années 1990. Jürgen Schmidhuber revendique ainsi le statut de véritable inventeur du *Deep Learning*, presque vingt ans avant tout le monde. Il crée avec l'un de ses étudiants, Sepp Hochreiter, les *Recurrent Neural Network* et les *Long Short-Term Memory* (LSTM), pour fournir de la « mémoire » à l'IA. Schmidhuber lui-même estime cependant que la toute première description du *Deep Learning* a été donnée en 1965 par deux chercheurs ukrainiens, Alexey Ivakhenko et Valentin Lapa.

Pour autant, ces méthodes mettront encore vingt ans avant de connaître leur premier succès. Les rares chercheurs à s'y intéresser et à les utiliser ne parviennent pas à faire publier leurs études et n'obtiennent que peu de soutien institutionnel. Parmi eux, les parrains du *Deep Learning*, Geoff Hinton, Yann Le Cun et le Canadien Yoshua Bengio, persistent avec le soutien du Canadian Institute for Advanced Research (CIFAR). Ils obtiendront tous trois le prix Turing en 2019, l'équivalent du Nobel en informatique. Seulement, les résultats obtenus à l'époque par les réseaux neuronaux sont comparables à ceux obtenus par d'autres méthodes, par exemple grâce aux machines à vecteur de support (S*upport Vector Machine*, SVM). La raison pour laquelle les performances du *Deep Learning* ne sont pas meilleures que les autres est, d'après Yann Le Cun, le manque de données. Pour faire une bonne intelligence artificielle avec du *Deep Learning*, il faut en effet de bonnes données (c'est-à-dire pas de données manquantes ou fausses) et en grande quantité.

En 2006, Fei-Fei Li, chercheuse en intelligence artificielle à Princeton, commence à travailler sur l'idée d'*ImageNet*, une base de données d'images

disponibles pour entraîner des algorithmes d'IA. En 2007, elle rencontre Christiane Fellbaum, l'une des créatrices de *WordNet*, pour discuter du projet. En se servant de cette base de données de mots et de ses caractéristiques, Li commence à construire *ImageNet*. Elle monte une équipe de chercheurs et utilise le service *Amazon Mechanical Turk* pour labelliser les images. *Mechanical Turk* est un service en ligne qui permet de *crowdsourcer* (faire appel à un très grand nombre de contributeurs *via* Internet) des tâches fastidieuses. Elle tire bien évidemment, et très ironiquement, son nom du fameux Turc mécanique. Progressivement, plus de 14 millions d'images seront labellisées et classées en plus de vingt mille catégories. À partir de 2010, le projet *ImageNet* organise une compétition annuelle appelée *ImageNet Large Scale Visual Recognition Challenge* (ILSVRC) au cours de laquelle différentes équipes d'ingénieurs et des chercheurs comparent les performances de leur IA en matière de reconnaissance d'image. Cette compétition sera à l'origine de nombreux progrès dans le domaine.

En 2011, Dan Ciresan, un chercheur postdoc de Schmidhuber, met au point le réseau neuronal profond *DanNet* et gagne toutes les compétitions mondiales, ainsi que les principales compétitions internationales de *Machine Learning* appliqué à la médecine : l'*International Symposium on Biomedical Imaging* (ISBI) pour la segmentation d'image cérébrale en 2012 et l'*International Conference on Pattern Recognition* (ICPR) la même année pour la détection de cancer du sein.

Toujours en 2012, à l'*European Conference on Computer Vision* (ECCV), le *Deep Learning* démontre sa suprématie. Alex Krizhevsky, un doctorant de Geoff Hinton, présente les résultats de son réseau neuronal profond appelé *AlexNet*. Alors que les meilleures équipes parviennent à un taux d'erreur de 27 %, Krizhevsky, plus ou moins inconnu à l'époque dans le domaine de la vision par ordinateur, rapporte un taux d'erreur de reconnaissance d'objets sur la base de photos *ImageNet* extrêmement bas de 17 %. C'est une authentique révolution. Dans la salle, l'auditoire est abasourdi. Krizhevsky n'a pas eu besoin de créer des règles pour décrire des images et ne possède pas non plus les connaissances habituelles des experts dans le domaine. Il a en revanche utilisé la puissance de calcul de plusieurs cartes graphiques (*Graphics Processing Unit*, GPU) pour entraîner un réseau neuronal profond à reconnaître les images, alors que la majorité des chercheurs ont recours aux processeurs CPU (*Central Processing Unit*) plus classiques, mais moins adaptés au calcul de *Deep Learning*. Fei-Fei Li, Yann Le Cun et tous les experts du *computer vision* sont stupéfaits. En une présentation, le jeune Alex Krizhevsky remet en cause dix ans de recherche en intelligence artificielle et en *natural object recognition*. Les chercheurs présents lui posent des questions méthodologiques qu'il ne comprend pas toujours. Yann Le Cun doit parfois répondre pour lui.

La course à l'armement en intelligence artificielle est donc officiellement lancée en 2012. À soixante-quatre ans, Geoff Hinton fonde une start-up

appelée DNN-Research avec ses deux étudiants (Ilya Sutskever et Alex Krizhevsky). Geoff reçoit une offre de rachat de douze millions de dollars de Baidu, géant numérique chinois, mais souhaite solliciter d'autres acheteurs. Son idée est donc de mettre aux enchères sa start-up afin d'être recruté par un géant du numérique. Geoff Hinton a de graves problèmes de dos et ne s'assoit plus depuis 2005. Il travaille et mange debout ou à genoux. Ne pouvant prendre l'avion pour cette raison, il emprunte donc un bus, dans lequel il peut s'allonger, afin de se rendre de Toronto à New York. À New York, il prend le train jusqu'à Truckee en Californie, traversant les États-Unis d'est en ouest. Il termine son trajet en taxi, allongé sur la banquette arrière, pour atteindre South Lake Tahoe (Californie), où aura lieu l'enchère. C'est dans la chambre 731 de l'Harrah and Harvey's, un hôtel-casino de la ville, que s'écrit pour partie l'avenir de l'IA. Hinton installe son ordinateur portable sur une poubelle retournée afin de pouvoir l'utiliser debout et lance les enchères. En parallèle, les plus grandes entreprises numériques du monde se préparent : Google (États-Unis), Microsoft (États-Unis), Baidu (Chine) et une petite start-up londonienne alors inconnue, DeepMind (cofondée par un élève de Schmidhuber). Les enchères se déroulent par mails. Chaque compétiteur envoie son offre à Geoff Hinton qui en informe ensuite les autres. Si aucune nouvelle offre n'est reçue au bout d'une heure, l'enchère est terminée.

L'adresse Gmail de Geoff Hinton inquiète un peu les concurrents qui craignent que Google ne lise ses mails pour prendre l'avantage. La start-up DeepMind se retire de l'enchère. Cela ne l'empêchera pas de devenir des années plus tard l'une des start-up les plus importantes dans le domaine, et même d'être rachetée par Google !

Rapidement, le montant du rachat de DNN-Research s'envole. Microsoft, Google et Baidu ont compris l'importance des travaux de Hinton et ce qui est en train de se passer. Microsoft se désengage pourtant. Ne restent plus que Google et Baidu. L'enchère atteint quarante-quatre millions de dollars. À la surprise générale, Hinton arrête l'enchère. Google acquiert DNN-Research et prend dès lors une avance considérable dans le domaine. Rétrospectivement, quarante-quatre millions de dollars peuvent sembler hors de prix. En réalité, comparé aux dix-neuf milliards de dollars déboursés par Facebook pour acheter WhatsApp, le montant est ridicule.

Après 2012, c'est l'emballement. Chaque mois, de nouveaux types de réseaux neuronaux profonds sont publiés avec toujours plus de couches et toujours plus de paramètres à optimiser (voir l'annexe « Bases techniques du *Deep Learning* »). Le *Deep Learning* sera une des principales voies de développement de l'intelligence artificielle pour la décennie à venir. Les grands acteurs industriels numériques sortent les uns après les autres leur propre bibliothèque de développement pour tenter d'asseoir leur suprématie : *Tensorflow* pour Google, *PyTorch* pour Facebook, *MXNet* pour Amazon, *Microsoft Cognitive Toolkit* pour Microsoft. Ces bibliothèques, disponibles en accès libre, sont comme des briques logicielles à utiliser les unes avec les

autres pour créer des IA que l'on entraîne sur de nouvelles données. Bien loin du Perceptron Mark 1, la machine physique laisse place à une virtualisation complète ouvrant des possibilités de calcul extraordinaires, notamment grâce à la puissance des GPU. Entre 2012 et 2020, le *Deep Learning* fait les gros titres des journaux scientifiques et de la presse généraliste. L'IA s'immisce alors dans tous les domaines de la vie quotidienne : elle reconnaît nos visages et nous identifie sur les photos de nos réseaux sociaux, elle gère nos contacts, nos mails, nos applications et notre calendrier sur notre téléphone, elle lit notre écriture pour encaisser nos chèques, la liste est infinie.

En 2016, *AlphaGo*, IA créée par DeepMind, bat le champion du monde du jeu de go, Lee Sedol. Trois ans plus tard, *AlphaStar*, une IA créée encore une fois par DeepMind, écrase les meilleurs joueurs mondiaux de Starcraft 2, un jeu vidéo de stratégie militaire. Au début des années 2020, certains experts pensent déjà que l'augmentation du nombre de paramètres et du nombre de couches des réseaux neuronaux profonds, ou même de la puissance de calcul des GPU, ne permettra pas de dépasser le plateau actuel de performance de l'IA. Il ne suffirait pas d'avoir des réseaux neuronaux profonds pour créer l'intelligence : ainsi, un nouveau cycle pourrait s'ouvrir, avec le retour d'une IA symbolique qui reviendrait pour s'hybrider avec l'IA connexionniste du *Deep Learning*.

Dans le domaine médical, l'utilisation de l'IA intéresse les chercheurs et les médecins dès la fin des années 1960. En 1970, le docteur William Schwartz publie dans le *New England Journal of Medicine* un article visionnaire intitulé : *Medicine and the Computer : The Promise and Problems of Change* (La médecine et l'ordinateur : les promesses et les problèmes d'une révolution). Il explique notamment que les sciences informatiques augmenteront et, dans certains cas, remplaceront les fonctions intellectuelles du médecin ; ce remplacement aura de multiples conséquences sur les plans psychologique, organisationnel, légal, économique et technique. Schwartz développe déjà l'idée que les interactions entre médecins, informaticiens et gestionnaires seront très importantes pour le bon développement de ces nouvelles techniques. Presque cinquante ans plus tard, ces thématiques sont plus que jamais d'actualité.

Les premiers résultats en intelligence artificielle appliquée à la médecine sont obtenus avec les systèmes DENDRAL (*Dendritic Algorithm*), un système expert en chimie organique, et SUMEX-AIM (*Stanford University Medical Experimental Computer – Artificial Intelligence in Medicine*), un système d'IA médicale en réseau, mis en place à l'université de Stanford dans le milieu des années 1970. Il s'agit de ressources informatiques partagées au niveau national, mettant à disposition des applications en biomédecine. Nous sommes encore loin d'une véritable intelligence artificielle.

En 1975, les principaux chercheurs dans le domaine échangent leurs idées sur les applications de l'intelligence artificielle à des thématiques

biomédicales dans une série d'ateliers financés par le National Institute of Health américain à l'université de Rutgers. L'un des pionniers, le médecin et informaticien canadien Edward H. Shortliffe, étudie les interactions patient-médecin médiées par ordinateur à l'université de Harvard. Il poursuit ses recherches à Stanford et rencontre Bruce Buchanan, informaticien qui travaillait déjà sur DENDRAL. Ensemble, ils mettent au point un système expert fondé sur des règles, capable de résoudre certains problèmes cliniques, en collaboration avec le biochimiste Stanley Cohen. Ces collaborations mènent à la création du système MYCIN qui guide alors l'utilisation des antibiotiques en maladie infectieuse. Ce système a notamment la particularité de fournir, en plus de sa recommandation, le niveau de « confiance » que l'on peut accorder à son conseil, ce qui permet d'augmenter son adoption par les médecins.

Le second système, *Causal-Associational Network* (CASNET), apparu dans les années 1970, est dédié au diagnostic et au traitement du glaucome ophtalmique et développé à l'université de Rutgers (États-Unis d'Amérique). À la même époque, émerge un troisième système, INTERNIST, fruit de la collaboration entre le biologiste Jack Myers et le chercheur en intelligence artificielle Harry Pople de l'université de Pittsburgh. Alors que MYCIN et CASNET sont dédiés à des maladies bien précises, INTERNIST est capable de couvrir l'ensemble de la médecine interne, un champ bien plus vaste. Malgré leurs succès théoriques, ces systèmes peinent cependant à s'imposer dans la pratique clinique des médecins.

De la fin des années 1970 jusqu'au début des années 2000, les principaux développements en informatique médicale sont consacrés à la création de systèmes utilisés en imagerie afin d'assister la classification et la lecture des examens (*Computer-Aided Diagnosis*, CAD). Ces systèmes, fondés initialement sur des règles simples, sont bouleversés par les récentes évolutions techniques du *Machine Learning*, notamment du *Deep Learning* comme nous le verrons dans les chapitres suivants. Cependant, vous l'aurez compris, à l'aube du XXI^e^ siècle l'« intelligence artificielle » est capable d'accomplir une tâche bien précise, mais pas davantage. En réalité, nous avons affaire à une multitude de petites intelligences artificielles limitées. Certains experts considèrent donc que l'emploi du terme *intelligence artificielle* est abusif et qu'il faut attendre l'apparition d'une intelligence artificielle générale, c'est-à-dire capable de réaliser un grand nombre de tâches différentes, avant de pouvoir utiliser cette expression à bon escient.

Résoudre le mystère de l'intelligence n'est pas simple. Plusieurs chercheurs estiment qu'il faudrait renoncer à simplement copier la nature, le cerveau avec ses réseaux de neurones. Car, un peu comme Clément Ader qui, en 1897 avec son *Avion III*, pensait voler en s'inspirant des chauves-souris, notre tendance à rechercher le mimétisme, à reproduire ce que nous croyons être l'intelligence, est sans doute la raison pour laquelle nous ne parvenons pas à la reproduire artificiellement : les avions d'aujourd'hui n'ont pas grand-

chose à voir avec des chauves-souris.

II

Utiliser l’intelligence artificielle pour mieux soigner

1

De la Silicon Valley et des robots

Après ma thèse, j'ai la chance extraordinaire de poursuivre mes recherches à Stanford où je débarque fin novembre 2019 avec ma femme et ma fille. L'histoire de l'intelligence artificielle est très intimement liée à celle de cette université de la Silicon Valley qui a formé un grand nombre d'équipes et accueille les plus grands chercheurs, même si elle n'a évidemment pas l'exclusivité de la recherche en IA.

La Leland Stanford University est une université privée fondée par Leland et Jane Stanford en 1891. Leland était un magnat des chemins de fer, gouverneur et sénateur de Californie. Le couple n'eut qu'un enfant, appelé Leland Stanford Junior, emporté par la fièvre typhoïde en 1884, peu avant son seizième anniversaire. Le couple décida de fonder et de financer une université pour les enfants de Californie sur l'ancien site de leur ferme à Palo Alto. Pour cette raison, Stanford University est souvent surnommée « The Farm ». Dès mon arrivée, je suis impressionné par la taille du campus qui s'étend sur trente-deux kilomètres carrés. Son architecture de style néoroman rappelle celle des missions qui jalonnaient El Camino Real lorsque la Californie était espagnole. Ses bâtiments sont connus dans le monde entier. Par exemple, la Stanford Memorial Church, construite par Jane Stanford au décès de son mari, ou la Hoover Tower, achevée en 1941, appartenant à la Hoover Institution, centre de recherche fondé par Herbert Hoover, ancien président des États-Unis. L'université possède une collection de plus de huit millions de livres répartis dans dix-neuf bibliothèques. Le campus abrite plusieurs musées, dont une importante collection de sculptures de Rodin. L'Industrial Park, créé en 1951, accueille des entreprises innovantes. Dans les structures de recherche, on retrouve le Stanford Linear Accelerator Laboratory, un laboratoire de physique des particules, ou la *Dish*, une antenne de radio de quarante-six mètres qui a notamment servi pour guider certains vaisseaux de la NASA tels que les missions *Voyager*.

Considérée comme la première ou la deuxième meilleure université du monde selon les classements, Stanford compte parmi ses professeurs vingt prix Nobel. Son développement est étroitement lié à celui des nouvelles technologies : elle a participé à l'élaboration d'Internet, et accueilli sur ses bancs William Hewlett et David Packard, fondateurs de Hewlett-Packard, Sergey Brin et Larry Page, cofondateurs de Google, ou encore Reed Hastings, PDG de Netflix.

À l'origine, la vallée de Santa Clara était connue pour ses nombreux

vergers. Le nom de Silicon Valley lui fut donné en 1971 par Don Hoefler, un journaliste local. Il fait référence au matériau de base des puces électroniques, le silicium. En 1956, William Shockley, physicien américain et lauréat du prix Nobel pour l'invention du transistor en 1948, fut recruté par Beckman Instruments pour créer son laboratoire à Mountain View. Pour ce faire, il engagea de jeunes diplômés qui formèrent le cœur de son équipe. Un an plus tard, huit de ces jeunes, que l'on appelle maintenant les « *Traitorous Eight* », les huit traîtres, quittèrent son laboratoire pour créer juste à côté, à San José, la Fairchild Semiconductor. Ils reprochaient à Shockley de se montrer trop rigide et entendaient créer une entreprise à l'ambiance plus détendue, à l'origine de l'esprit de la Silicon Valley. Fairchild fut l'une des entreprises pionnières dans le domaine des circuits intégrés et la pépinière de nombreux talents. C'est en quelque sorte la grand-mère de toutes les start-up qui s'affranchissent maintenant des codes vestimentaires costume-cravate.

En 1963 Andrew Grove, Gordon Moore et Robert Noyce, qui travaillaient chez Fairchild, fondèrent Intel. En 1969, d'autres employés de Fairchild créerent AMD. Ces deux entreprises sont maintenant leaders dans le domaine des microprocesseurs.

Après un passage à vide dans les années 1980, l'émergence du Web et celle d'Internet réaffirment le rôle prépondérant de la Silicon Valley dans les hautes technologies. Stanford accueille de nos jours de très nombreuses entreprises hardware (telles que Intel, AMD, Nvidia), software (Google, Facebook, Apple), mais aussi médical et biotech (Varian, Verily, Genentech).

Très vite, je constate que les robots sont très présents dans la région, du robot domestique au véhicule autonome, en passant par les gardiens de parking ! L'un des premiers robots grand public a été commercialisé en 1999 par Sony : le chien AIBO, pour *Artificial Intelligent Robot*. J'ai toujours rêvé d'en posséder un (non parce qu'il ressemblait à un chien mais parce que je voulais un robot !). Les *Aibos* ont été développés par le Computer Science Laboratory de Sony à partir de 1994. Une première version a été commercialisée à 5 000 exemplaires et une seconde version plus perfectionnée est sortie plus récemment en 2018. Malheureusement elle n'est disponible qu'au Japon et aux États-Unis. Mon séjour est donc l'occasion parfaite de me procurer le robot. Et je bénéficie de l'excuse de ma petite fille de 3 ans, parfait alibi pour justifier l'achat auprès de ma femme ! Je commence donc mes recherches par eBay. Parmi les nombreuses annonces, plus ou moins sérieuses et plus ou moins chères, une offre attire mon attention : le vendeur se trouve à Palo Alto, à dix minutes de chez moi. Je décide donc de placer une enchère.

Au bout d'une semaine, je reçois une notification sur mon téléphone : j'ai remporté l'enchère de l'*Aibo* ! Vient le moment de prendre contact avec le vendeur. Je commence alors à douter et me dis que l'affaire est trop belle pour être vraie, j'ai probablement été arnaqué. Le soir suivant, en sortant du laboratoire, je file en voiture sur Alma Street récupérer le robot. Le vendeur a

à peu près le même âge que moi. Nous échangeons quelques banalités et il ouvre rapidement le coffre de sa Tesla Model X pour me remettre le carton d'origine contenant l'*Aibo* en parfait état. Avant de me le remettre, un peu à regret me semble-t-il, il me demande si j'ai un intérêt particulier pour la robotique. Je lui réponds que je fais de la recherche appliquée en intelligence artificielle mais que c'est davantage par curiosité ainsi que pour ma petite fille que j'achète le robot. Nous nous séparons et je rentre rapidement chez moi. Après quasiment vingt ans, j'ai enfin un *Aibo*, adopté directement par ma fille. Ce chiot robotique est capable de marcher, de répondre aux ordres et de monter la garde. Il peut également être programmé pour accomplir de nouvelles tâches. Il restera avec nous pendant tout notre séjour californien et rentrera même avec nous à Paris.

Mais l'*Aibo* n'est pas la seule rencontre avec un robot à avoir marqué notre séjour dans la Silicon Valley. Un soir, alors que nous allons voir le dernier film de Quentin Tarantino au Century Cinema 16 de Mountain View, nous nous garons dans un parking sur le North Shoreline Boulevard, juste à côté du campus de Google. À notre grande surprise, le parking est gardé par un robot équipé d'une caméra. Si vous avez vu le film de 1956 *La Planète interdite*, vous connaissez sans doute Robby le robot, véritable star du film. Le robot chargé de garder le parking du Century Cinema 16 de Mountain View est une sorte de Robby du futur, capable de vous rappeler à l'ordre si vous êtes mal garé, ou de surveiller votre voiture en votre absence. Il est commercialisé par la start-up Knightscope dont la mission est, je cite, « de faire des États-Unis d'Amérique le pays le plus sûr du monde » grâce à toute une gamme de robots de sécurité autonomes – tout un programme. Le modèle en question est un *K-5*, équipé d'un LIDAR (télédétection par laser), d'un sonar, d'un GPS et de caméra à 360 degrés capables de filmer vingt-quatre heures sur vingt-quatre. L'entreprise se vante d'une réduction de 46 % des crimes, d'une augmentation de 27 % des arrestations et d'une diminution de 68 % des procès-verbaux. Heureusement, il n'est pas encore armé ! Ses camarades sont très nombreux dans la Silicon Valley.

Autre exemple : sur le chemin du laboratoire, je croise quotidiennement les véhicules autonomes de la start-up Waymo, une filiale de Google. De très grands vans blancs, immédiatement reconnaissables à leur multitude de capteurs, qui peuvent servir de taxis dans la Bay Area et sont actuellement en test. La législation concernant le test de véhicules autonomes en Californie, bien qu'assez stricte, est beaucoup plus souple que celle de la France. Un grand nombre d'entreprises ont investi dans ce domaine, de Google à Uber, en passant bien sûr par Apple et son projet *Titan*. L'une des plus avancées est sans nul doute Tesla qui propose déjà une option de conduite autonome complète sur autoroute. Plusieurs vidéos sur Internet montrent des conducteurs endormis au volant de leur *Tesla* filant à 70 miles/heure sur la 101, célèbre *freeway* qui relie le sud de la Bay Area et San Francisco.

Le domaine médical compte aussi quelques robots. Parmi eux, le robot

chirurgical *Da Vinci*, commercialisé par Intuitive Surgical, qui permet des interventions très peu invasives. Autorisé depuis 2000 aux États-Unis par la Food and Drug Administration (FDA), ses filtres logiciels permettent de supprimer les tremblements du chirurgien et d'exécuter des gestes plus précis. Le chirurgien s'installe à une console, visualise le champ opératoire grâce aux écrans vidéo et commande les bras du robot pour opérer le patient grâce à des manettes semblables aux joysticks des consoles de jeux vidéo. Il ne s'agit donc pas d'un robot autonome, mais l'ajout d'une IA pourrait un jour déboucher sur des traitements chirurgicaux entièrement automatisés.

Plus autonome, le *CyberKnife* est un robot de radiothérapie commercialisé au début des années 2000 par Accuray. Inventé par John Adler, un professeur de neurochirurgie de Stanford, cet appareil permet de traiter des tumeurs dans l'ensemble du corps avec une très grande précision, en délivrant des photons de haute énergie. Il comporte un accélérateur de particules miniaturisé et monté au bout d'un bras robotique qui peut le guider dans toutes les dimensions de l'espace. Une fois définie la zone à traiter, la machine agit de façon totalement autonome en s'adaptant à la position du patient et à ses mouvements respiratoires, sans aucune intervention humaine. Elle possède un système d'imagerie embarqué (radiographie et caméra infrarouge) qui lui permet à la fois de visualiser précisément le patient et la tumeur et d'administrer le traitement. Bien que moins connu que le *Da Vinci*, le *CyberKnife* est le premier et le seul robot médical à administrer le traitement de façon autonome. Ce type de radiothérapie, appelée stéréotaxique, permet de guérir certains cancers inopérables pour lesquels nous ne disposions d'aucune option thérapeutique il y a encore quelques années.

Dans les chapitres suivants, je vous propose de discuter des applications de l'IA en médecine dans le dépistage, le diagnostic, la prédiction et le traitement des maladies. Nous verrons également comment l'IA peut être utilisée pour prévenir leur apparition et comment elle peut nous aider dans notre bien-être au quotidien.

2

Dépister grâce à l'IA

L'épidémiologie est l'étude des rapports entre les maladies et les facteurs qui sont susceptibles de les causer ou d'influencer leur fréquence, leur distribution dans l'espace et leur évolution. Elle revêt une importance primordiale, notamment dans les politiques de santé publique, en permettant par exemple de détecter ou de prévenir l'apparition de nouvelles maladies ou la résurgence d'anciennes. L'épidémiologie en tant que champ d'étude propre n'a été reconnue que relativement récemment, en 1854. Elle est pourtant l'un des piliers de la médecine en général et de la médecine préventive en particulier. L'épidémiologie a toujours eu recours aux méthodes statistiques et aux théories des probabilités. En ce sens, ce que l'on appelle aujourd'hui les *big data* y sont utilisés depuis très longtemps. L'intelligence artificielle peut avoir de nombreuses implications dans les études analytiques qui cherchent à définir les déterminants environnementaux d'une maladie. Notamment quand il s'agit de trouver des corrélations entre certains comportements ou certaines caractéristiques sociodémographiques et l'apparition de maladies comme l'obésité, le diabète ou le cancer.

Les causes de l'augmentation actuelle du nombre de cancers dans le monde s'expliquent en partie par les modifications de nos comportements et de notre environnement et bien sûr par le vieillissement de la population. Dans l'objectif d'adapter les politiques de prévention et les moyens de santé, les données épidémiologiques sur la prévalence du cancer sont primordiales. Cependant, les registres épidémiologiques actuellement disponibles sont incomplets et nous ne disposons d'aucune évaluation très précise pour de vastes régions. C'est pour cette raison que je me suis intéressé à cette thématique. Des équipes américaines avaient déjà publié des études montrant que l'on pouvait utiliser l'IA afin d'analyser des images satellite pour prédire la pauvreté au Nigeria ou déterminer la prévalence de l'obésité aux États-Unis, mais aucune ne s'était intéressée à la prédiction de la prévalence du cancer, alors que l'on sait que cette prévalence dépend de l'environnement. De réunions de laboratoire en séances de brainstorming, nous sommes parvenus à identifier les sources de données que nous pourrions utiliser pour ce projet : nous devions commencer par télécharger des images satellitaires de Google Maps.

Au cours de mon séjour aux États-Unis, j'ai eu plusieurs fois l'occasion de visiter les bâtiments du *Googleplex*, siège social de l'entreprise emblématique de la Silicon Valley. Les principaux sont situés à Mountain View, ville où

l'histoire de la Silicon Valley a commencé en 1956. Google a racheté l'ancien campus de Silicon Graphics érigé entre 1994 et 1997 et s'y est installé à partir de 2003. Tout autour de ce complexe se trouvent des restaurants et des terrains de sport destinés aux employés. De nombreux vélos aux couleurs de l'entreprise sont mis à disposition pour aller d'un point à l'autre. Il est très étonnant de s'apercevoir que l'on peut pénétrer sur le campus sans grande difficulté. L'endroit évoque fortement le siège de la start-up fictive Amaya de la série *DEVS* (Alex Garland, 2019), où une jeune pousse travaille secrètement au développement d'un ordinateur quantique au milieu des séquoias géants typiques de la région.

Par contraste, aucun signe extérieur ne permettait d'identifier les nouveaux bâtiments de Google Health à Palo Alto où je me suis rendu plusieurs fois pour des réunions inter-équipes centrées sur nos recherches et coordonner l'écriture d'un ouvrage sur l'intelligence artificielle en médecine. Fallait-il que la branche santé reste discrète ? L'un des principaux projets de Google Health était *Nightingale* : à partir des données de 50 millions de patients américains traités dans le réseau de cliniques privées *Ascension*, Google souhaitait créer des algorithmes prédictifs et aussi révolutionner le dossier médical informatisé, en intégrant un moteur de recherche convivial, intuitif et rapide, permettant d'explorer les dossiers médicaux des patients beaucoup plus facilement que ce qui existait. En novembre 2019, un lanceur d'alerte qui travaillait chez Google contacta le *Guardian* pour révéler que les données dont disposait Google incluaient les noms des patients et un très grand nombre de données personnelles[3]. Cette controverse obligea Google à s'expliquer et entacha durablement la réputation de la division. Quoi qu'il en soit, à l'été 2021, la division santé fut finalement dissoute, deux ans après sa création, et les employés assignés à d'autres projets. Google la relancera en 2022.

Pour en revenir à nos travaux sur les facteurs environnementaux du cancer, nous devions en parallèle des images Google Maps obtenir leurs coordonnées GPS puis en extraire un vecteur mathématique, c'est-à-dire un ensemble de valeurs numériques qui représentent correctement l'image, à un certain niveau d'abstraction. Nous avions également besoin de connaître la prévalence du cancer dans les zones géographiques dont nous avions les images. J'avais décidé de cibler sept des plus grandes villes américaines pour tester la méthode. Il fallait donc télécharger plus de 170 000 images de ces villes, ainsi que leur position GPS précise... Pour télécharger ces images, j'ai dû laisser tourner les stations de travail du laboratoire plusieurs jours durant.

Après avoir obtenu toutes les images, nous avons commencé à utiliser différents réseaux neuronaux pour extraire un vecteur mathématique de chaque image. Nous avons ainsi testé les capacités d'une dizaine de réseaux neuronaux classiques, tels que VGG-16 ou *ResNet*, ou plus modernes, comme *EfficientNet*, le tout dernier réseau de Google sorti à l'été 2019. Alors que ces réseaux sont utilisés d'habitude pour identifier des objets sur des images, nous nous en sommes servis pour extraire les caractéristiques numériques des

images satellitaires. La solution était de « décapiter », c'est-à-dire d'enlever la dernière couche du réseau qui fait habituellement le rapprochement entre le vecteur mathématique extrait de l'image et un objet. Durant cette étape, il était véritablement fascinant de voir en direct la progression des réseaux pour extraire les vecteurs de ces centaines de milliers d'images. À la fin, nous avons donc obtenu un peu plus de 170 000 vecteurs. Pour vérifier que ces vecteurs fournissaient une représentation fidèle du contenu de l'image, nous avons réalisé deux expériences : la première a consisté à prendre une image d'un environnement donné, à en extraire son vecteur et à demander ensuite aux réseaux de retrouver des images similaires parmi toutes les images, uniquement à partir du vecteur mathématique. Miracle, cela fonctionnait parfaitement : donnez un vecteur d'une zone pavillonnaire au réseau et il retrouvera et affichera toutes les zones pavillonnaires similaires. La seconde expérience a consisté à rassembler tous les vecteurs similaires par groupe puis à regarder si les images qui leur correspondent montrent des environnements proches. Là encore, l'expérience était très concluante et nous avons pu générer pour chaque ville des mosaïques où chaque type d'environnement était parfaitement groupé.

J'ai ensuite écrit avec l'aide de Hongyi Ren, un étudiant en master du laboratoire, un script capable d'interroger à très haute fréquente le site de la FCC (Federal Communications Commission) pour déterminer à quelle zone de recensement (*census tract*) appartenait chaque image, à partir du fichier de *geotagging* GPS. Les *census tracts* sont importants : ils correspondent à des zones de 4 000 à 6 000 habitants pour lesquels nous avions la prévalence précise du cancer, grâce aux registres épidémiologiques du Center for Disease Control (CDC) américain.

La dernière étape, la plus simple, a consisté à entraîner un modèle de type *ElasticNet* pour corréler la prévalence au vecteur. Cet entraînement a été réalisé sur 80 % des images, ville par ville, et les résultats vérifiés ensuite sur les 20 % des images restants. Au final, nous avons pu montrer que cette méthode fonctionnait bien et qu'elle permettait de prédire jusqu'à 65 % de la prévalence du cancer uniquement à partir des images satellite, sans aucune autre information. Pour comprendre comment marchait l'IA que nous avions créée, nous avons vérifié la corrélation entre l'environnement et d'autres déterminants connus du risque de cancer, comme l'âge, l'obésité, le diabète, le manque d'activité physique, et nous avons pu constater qu'il existait un lien statistique très fort[4]. En somme, l'environnement dans lequel nous vivons est un excellent reflet de nos comportements et de nos facteurs de risque de développer des maladies. Pour terminer, nous avons voulu vérifier si l'IA ainsi créée était capable de prédire la prévalence du cancer sur une zone géographique qu'elle ne connaissait pas du tout. Nous avons donc pris le modèle créé pour Los Angeles et vérifié ses résultats sur une autre ville comme Phoenix par exemple, et ainsi de suite pour les sept villes dont nous disposions. Comme cela était attendu, nous avons pu montrer que, si deux

villes ont des environnements proches, le modèle de l'une peut être utilisé pour une autre. Notre IA pourrait donc être utilisée pour déterminer la prévalence du cancer dans une zone où elle est inconnue. Les applications potentielles sont nombreuses pour mieux comprendre le rôle de l'environnement dans le risque de développer la maladie et guider les politiques de santé publique. Afin de pouvoir guider les programmes de dépistage du cancer ou les politiques de prévention (contre le tabagisme, l'alcoolisme ou le manque d'activité physique), il est d'abord nécessaire de savoir où le taux de cancer est important afin de cibler ses interventions pour qu'elles soient plus efficaces. Obtenir des cartographies précises de la prévalence du cancer coûte extrêmement cher. Par exemple, les États-Unis ont dépensé 22 millions de dollars pour recueillir ce type de données couvrant approximativement 20 % de la population. En France, ces données ne sont malheureusement pas disponibles à une résolution spatiale suffisante (6 000 à 8 000 habitants). Pourtant cette approche pourrait être utilisée pour obtenir les taux de cancer dans les régions où l'on ne le connaît pas, à moindre coût.

Les applications de l'IA en épidémiologie ne s'arrêtent pas au cancer : cette méthodologie pourrait être appliquée au dépistage d'autres types de maladies fortement corrélées à l'environnement. On pense évidemment aux maladies infectieuses telles que la Covid-19. En utilisant la même méthode, le *Deep Learning* nous aiderait à comprendre l'évolution d'une pandémie, afin de l'anticiper et d'adapter les mesures de contrôle sanitaire. Les données sont susceptibles d'être enrichies par les recherches de symptômes sur Internet (car on sait que les personnes malades les *googlisent*), les informations de flux de voyage par avion ou par train, ou encore les conditions météorologiques. En 2019, une équipe de l'université de Tokyo a ainsi utilisé les données de température de la mer et de pluviométrie afin de prédire la distribution spatiotemporelle de la dengue à Taïwan. En partant du principe que l'on peut utiliser l'intelligence artificielle pour modéliser la distribution des nids de moustiques, on peut aussi lutter contre les maladies infectieuses comme le virus Zika, la dengue ou le chikungunya.

Au-delà de ces données du monde réel, il est également possible d'utiliser les données d'Internet pour entraîner une IA en épidémiologie. Les systèmes de surveillance Internet sont classés en trois grandes catégories : surveillance modérée, surveillance partiellement modérée et surveillance automatisée. Parmi eux, le réseau intelligent de santé publique globale (*Global Public Health Intelligence Network*) a été développé par le gouvernement canadien. Il analyse environ trois mille dépêches d'actualité par jour. En Europe, nous disposons d'un système entièrement automatisé, appelé MedISys (*Medical Information System*) mis en place par la Commission européenne en 2004. Ce système analyse vingt-quatre heures sur vingt-quatre les articles et les dépêches sur 2 200 sites et cinquante langues différentes. Chaque jour, quelque 90 000 articles sont traités pour surveiller l'apparition de maladies infectieuses, le bioterrorisme et les menaces chimiques, biologiques,

radiologiques et nucléaires. Plus récemment, en 2019, le système SENTINEL a été créé afin d'exploiter les techniques d'intelligence artificielle et monitorer en temps réel les réseaux sociaux pour l'apparition d'épidémie. De façon globale, le système surveille l'apparition de mots-clés liés à certains symptômes et les quantifie afin de détecter une éventuelle épidémie.

L'une des sources de données les plus intéressantes dans ce contexte est évidemment les réseaux sociaux. Parmi eux, Twitter a démontré son utilité dans la surveillance épidémiologique de maladies infectieuses, que ce soit l'épidémie d'hantavirus au Chili ou les épidémies de grippe.

Cependant, même avec la meilleure des IA, nous aurons toujours besoin d'épidémiologistes, de virologistes et d'infectiologues sur le terrain pour dépister et combattre la maladie en première ligne. Ainsi, comme nous le verrons pour les autres étapes du parcours de soins, l'intelligence artificielle ne remplace pas les métiers déjà existants mais vient se positionner en outil supplémentaire qui enrichit la pratique et la rend plus efficace.

3

Diagnostiquer et prévoir grâce à l'IA

Après le dépistage, le diagnostic en médecine correspond à l'identification de la cause ou de la nature d'une maladie à partir de ses signes cliniques ou paracliniques. Il s'agit concrètement d'une tâche de corrélation ou de perception pure et c'est donc tout naturellement que l'intelligence artificielle excelle dans ce champ. L'intérêt pour les techniques de *Deep Learning* s'est très significativement accentué depuis une dizaine d'années. L'IA a en effet un très grand potentiel dans la reconnaissance et la localisation d'une anomalie sur une image. Elle peut être utilisée sur l'ensemble des sites anatomiques : la tête, le thorax, l'abdomen ou le pelvis.

L'ophtalmologie est l'une des premières spécialités où l'IA a été utilisée, plus particulièrement sur l'analyse du fond d'œil et des rétinopathies. L'étude la plus connue a été publiée dans le *Journal of American Medical Association* (*JAMA*) en 2016 par une équipe de Google[5]. Les chercheurs ont entraîné un réseau neuronal convolutif (CNN) avec les données de plus de 228 000 images rétiniennes afin de diagnostiquer une rétinopathie diabétique. Les performances de l'IA ont ensuite été comparées à celles de sept ophtalmologues américains experts sur un peu plus de 11 000 nouvelles images de rétines. Le réseau neuronal a fourni de meilleures performances diagnostiques que le groupe d'ophtalmologues.

La cancérologie, de par son aspect multidisciplinaire, et à cause de la très grande quantité de caractéristiques nécessaires à la définition d'un diagnostic et d'une stratégie thérapeutique, est sans doute l'une des disciplines qui seront les plus bouleversées par l'intelligence artificielle, notamment pour la personnalisation des soins. De nombreuses bases de données sont développées dans différents pays : à Paris, un projet développé par le programme CARPEM (*Cancer Research for Personalized Medicine*) rassemble notamment les dossiers médicaux d'hôpitaux de l'Assistance publique (Hôpital européen Georges-Pompidou et hôpital Cochin). Ce projet, auquel je participe, réunit trois cent cinquante chercheurs et médecins dédiés aux soins et à la recherche sur le cancer regroupés en vingt et une équipes de recherche, onze équipes cliniques, et soutenus par huit plateformes. Leur domaine de compétences couvre l'immunologie, la génomique, le métabolisme, la biologie cellulaire, la bio-informatique, les mathématiques, les biostatistiques, l'informatique médicale, la microfluidique, l'épidémiologie et l'éthique. En

Californie, une base de données appelée *OncoSHARE* collecte les informations du registre des cancers de Californie, des hôpitaux de l'université de Stanford et d'autres centres de soins. Aux États-Unis toujours, à une échelle nationale, l'American Society of Clinical Oncology (ASCO), la Société américaine d'oncologie clinique, a créé la base *CancerLinQ*, qui rassemble les données de santé de nombreux patients mais aussi des données épidémiologiques. Les exemples sont nombreux. Le but de ces bases de données est de prendre en compte la grande hétérogénéité des patients atteints de cancer, afin de développer de nouvelles stratégies thérapeutiques, sans avoir nécessairement besoin de refaire de nouveaux essais cliniques, longs et coûteux. Chaque patient et chaque cancer étant différent (antécédents, anomalies génétiques, stade de la maladie, sensibilité aux traitements), il est nécessaire de disposer de ces informations précises pour créer des IA performantes. L'administration américaine qui régule la mise sur le marché des nouveaux médicaments, la FDA, a bien compris l'intérêt de ce type de bases et a publié un guide de bonnes pratiques pour la création de *Real-World Evidence* (RWE), de nouvelles preuves fondées sur les données réelles.

En cancérologie, l'une des grandes difficultés que l'on peut rencontrer concerne le traitement du cancer de la prostate. Ce cancer, très fréquent, est une maladie d'évolution très lente. Pour cette raison, son dépistage par dosage sanguin du *Prostate Specific Antigen* (PSA) n'est pas systématiquement recommandé, particulièrement chez les patients âgées ou ayant d'autres maladies graves. Différentes études ont montré que ce test ne permettait pas toujours d'améliorer la survie des patients, mais entraînait au contraire un surdiagnostic et donc un surtraitement. Dans ce contexte, il semblerait donc pertinent d'utiliser les capacités prédictives d'une intelligence artificielle afin de distinguer les patients qui pourraient potentiellement bénéficier d'un traitement, de ceux pour qui ce serait inutile. C'est à partir de ce constat que j'ai mené avec l'équipe de Stanford un projet afin de créer un algorithme capable de prédire la survie de patients dix ans après le diagnostic d'un cancer de la prostate. Je prends donc cet exemple afin de vous décrire les différentes étapes de la création et les résultats obtenus par l'IA[6].

Pour créer cette IA, nous avons pu obtenir les données d'un essai clinique prospectif du National Cancer Institute aux États-Unis (essai *Prostate Lung Colon Ovary*, NCT 00002540). L'étude incluait 76 693 patients, et nous avons sélectionné les 8 776 patients qui se sont vu diagnostiquer un cancer de prostate durant leur suivi. Nous avons utilisé un modèle *XGBoost* (*Gradient-Boosted Tree*) afin de déterminer les patients à risque de décéder de leur cancer ou d'une autre cause, dix ans après le diagnostic. Le *Gradient boosting* est une technique de *Machine Learning* qui produit un modèle sous la forme d'une addition de sous-modèles prédictifs, dans notre cas des arbres décisionnels (*decision trees*).

La cohorte a été divisée en base d'entraînement et de validation. Les modèles obtenus ont d'excellentes performances discriminatives, qui sont

même meilleures pour prédire le risque de décéder du cancer de la prostate ou d'une autre cause que le test de dépistage du cancer de prostate (PSA et toucher rectal annuels).

Ce modèle permet donc de guider la prise de décision thérapeutique : en sachant à l'avance quel patient est à risque de décéder d'un cancer de la prostate ou d'une autre cause, on peut décider si traiter le cancer a un intérêt ou non. Cependant, sans possibilité de comprendre pourquoi le modèle fournit cette prédiction, il est difficile de lui faire confiance. Nous avons donc ajouté la possibilité de visualiser les caractéristiques qui sont à l'origine de la prédiction en utilisant les valeurs de Shapley. Cette méthode est issue de la théorie du jeu : elle évalue la contribution de chaque variable dans une prédiction donnée. La *Shapley value* est la moyenne pondérée de la contribution de la variable d'intérêt pour chaque combinaison de variables. Cela permet non seulement une quantification de l'importance de chaque variable sur l'ensemble de l'échantillon, mais aussi pour un sous-groupe de patients ou un seul patient. Les modèles ont ensuite été déployés en ligne[7] afin de permettre aux médecins de les utiliser pour guider la décision thérapeutique. En répondant à une série de trente questions simples, les modèles permettent d'obtenir la probabilité de décès sous dix ans et de visualiser les raisons motivant cette prédiction.

Au-delà des dossiers médicaux, l'analyse d'images par intelligence artificielle a la capacité de révolutionner de nombreux aspects de la cancérologie ; parmi ces techniques, on peut parler de la *radiomics*, qui consiste à ne plus considérer que les images contiennent simplement des informations morphologiques ou métaboliques, mais que l'on peut aussi en extraire des statistiques, invisibles à l'œil nu. Parmi elles, des descripteurs numériques de la taille et de la forme d'une tumeur ou de sa texture. L'utilisation d'un algorithme informatique révèle des informations échappant à la lecture humaine traditionnelle. On pourrait par exemple analyser un scanner ou une IRM automatiquement et donner les chances de guérison d'un cancer. C'est ce qu'a réalisé une équipe de chercheurs de l'Institut Gustave-Roussy qui a utilisé les images de cent trente-cinq patients afin de prédire l'efficacité de l'immunothérapie[8]. L'algorithme a ainsi appris ce dont le radiologue est incapable : il a reconnu, à partir des niveaux de gris des scanners, les textures caractéristiques des tumeurs sur lesquelles l'immunothérapie fonctionnera. Dans le domaine de l'imagerie appliquée à l'oncologie, les autres progrès potentiels liés à l'intelligence artificielle comprennent également la possibilité d'améliorer la qualité des images réalisées ou d'identifier automatiquement les organes figurant sur une imagerie 3D.

La radiologie est une des spécialités où l'IA a le plus grand impact. 4,4 millions de radiographies du thorax sont effectuées en France chaque année. Un système fondé sur l'IA capable de réaliser une lecture automatisée de ce

type d'examen représenterait un gain significatif de temps et potentiellement d'efficacité. Afin de créer ce genre de système, il est une fois de plus nécessaire de partir des données, et donc d'obtenir un très grand nombre de radiographies. Les clichés des systèmes d'imagerie des hôpitaux, les PACS (*Picture Archiving and Communication System*) sont de plus en plus utilisés. Cependant les radiographies seules ne suffisent pas, il faut aussi récupérer le diagnostic correspondant et, dans le meilleur des cas, la zone suspecte au sein de l'image. Il existe plusieurs bases de données disponibles comprenant entre 100 et 200 000 radiographies chacune (*CXR14*, *CheXpert*, *MIMIC-CXR* et *PadChest*). L'une des études les plus importantes dans le domaine a été menée par une équipe de l'université de Stanford et publiée en 2017.

Pour ce projet, les chercheurs ont créé un algorithme appelé *CheXNet*, correspondant à un réseau neuronal profond de 121 couches[9]. Ce réseau a été entraîné sur la base *CXR14*, comprenant un peu plus de 100 000 radiographies de thorax de face. Son but : être capable de diagnostiquer quatorze pathologies thoraciques comprenant l'infection pulmonaire, la pleurésie, le pneumothorax, ou l'existence d'une masse tumorale. Dans cette étude, les chercheurs ont comparé les performances de leur réseau aux performances de quatre radiologues experts et montré que, dans la majorité des cas, l'intelligence artificielle obtenait de meilleures performances. L'originalité de *CheXNet* tient aussi à sa capacité à fournir la carte de chaleur qui permet de visualiser la zone suspecte directement sur la radiographie. Ce type d'approche permet non seulement de comprendre la prédiction réalisée par l'intelligence artificielle mais guide également le radiologue qui peut ainsi gagner du temps et petit à petit avoir confiance dans les résultats fournis par l'algorithme.

Ce principe a récemment été repris dans des études évaluant la capacité d'une intelligence artificielle à diagnostiquer une pneumopathie Covid-19 sur des radiographies de thorax[10].

Au-delà de l'analyse bidimensionnelle, l'analyse d'images scanographiques en trois dimensions pourrait être un domaine d'application de l'intelligence artificielle encore plus intéressant.

Chaque année, le cancer du poumon est à l'origine d'un peu plus de 31 000 décès en France, ce qui en fait la première cause de mortalité par cancer dans notre pays. Aujourd'hui, le pronostic de ce type de cancer est encore sombre : seuls 17 % des patients sont en vie cinq ans après le diagnostic. En parallèle, le dépistage par scanner faible dose a montré son efficacité en entraînant une réduction de la mortalité de 20 à 40 %. Ce dépistage est actuellement limité par la disponibilité des machines pour réaliser des scanners, mais aussi des radiologues pour les interpréter. En partant de ce constat, une équipe de Google a créé en 2019 une intelligence artificielle capable de détecter une tumeur pulmonaire et même de prédire l'apparition d'un cancer du poumon. Pour ce faire, ils ont utilisé un peu plus de 14 800 cas (avec 42 290 scanners thoraciques) issus d'un essai clinique

américain de dépistage national du cancer du poumon (*National Lung Screening Trial*). Le système est constitué de quatre composants qui s'appuient sur la bibliothèque de *Deep Learning* créée par Google, *TensorFlow* (une des plus utilisées dans le monde) :

— segmentation automatique des poumons ;
— détection automatique de nodules tumoraux ;
— prédiction du risque d'apparition de cancer à un an ;
— synthèse du risque de cancer sur le scanner actuel et un an après.

L'équipe a ensuite validé son algorithme sur plus de 6 700 nouveaux cas (qui n'étaient pas connus par l'intelligence artificielle). Les performances obtenues sont excellentes, avec un score AUC (*area under the curve* – aire sous la courbe) de 94,4 %. Les membres de l'équipe ont enfin comparé les performances de leur IA à celles d'un groupe de six radiologues experts et montré la supériorité de leur approche entièrement automatisée[11]. Cette étude préfigure le type de systèmes qui pourrait être déployé à l'avenir dans les hôpitaux et les cabinets de radiologie. L'intelligence artificielle permet non seulement de détecter la présence d'un cancer sur l'examen qui vient d'être réalisé, mais prédit aussi le risque d'en développer un plus tard. En ce sens, il s'agit d'un algorithme capable de faire ce que savent faire les humains (détecter un cancer), mais aussi capable de faire ce que ne savent pas faire les humains (prédire l'apparition d'un cancer).

En utilisant ce type d'approche, nous favoriserions l'accès à un programme national de dépistage de cancer du poumon. Au-delà de la disponibilité des machines et des radiologues, l'IA pourrait diminuer le coût de ce type de programme et surtout réduire le nombre de décès liés au cancer du poumon en réalisant un diagnostic plus précoce permettant un traitement curatif. L'implication directe du déploiement de ce genre de méthodes est donc de sauver des vies à très grande échelle. Cependant, avant de déployer ce type d'intelligences artificielles, nous aurons besoin d'une validation prospective rigoureuse.

Dans le domaine de la radiologie abdominale, les autres applications du *Deep Learning* comportent notamment la détection automatique de cancer du foie ou du pancréas et la prédiction de son pronostic. Mais, plus généralement, l'une des voies les plus prometteuses concerne ce que l'on appelle l'analyse universelle de lésion, le *Universal Lesion Analysis* (ULA). Dans cette approche, on entraîne une intelligence artificielle à analyser un scanner de l'ensemble du corps d'un patient, afin de détecter par exemple un nodule pulmonaire, un ganglion métastatique, un cancer du sein ou toute autre anomalie. Au-delà de la détection de l'anomalie, ce type d'approche corps entier offre la possibilité de prédire la survie par exemple cinq ans à l'avance. C'est ce qu'une équipe australienne a démontré dans une étude parue en 2016[12]. Les chercheurs ont ainsi créé une intelligence artificielle capable de déterminer la probabilité de décès à cinq ans, uniquement à partir d'un scanner. Les chercheurs ont pour cela sélectionné vingt-quatre patients

décédés en 2014 qui avaient eu un scanner thoracique dans les cinq années précédentes, âgés d'au moins 60 ans. Il est important de comprendre que ces scanners ne présentaient aucune anomalie évidente à l'œil humain. En parallèle, les chercheurs ont obtenu les scanners de vingt-quatre autres patients « contrôle », afin d'entraîner une intelligence artificielle par *Deep Learning* à reconnaître sur les scanners des patients décédés les anomalies éventuelles, invisibles à l'œil nu. Bien que cela puisse sembler contre-intuitif, ils sont parvenus à obtenir de bons résultats et à prédire de façon correcte la mortalité à cinq ans, à partir d'un scanner thoracique qui était *a priori* normal. On comprend dès lors la puissance potentielle de ce type d'approche mais aussi les éventuelles dérives susceptibles d'en découler. En effet, imaginez que vous alliez aux urgences, que vous passiez un scanner et que l'on vous annonce à cette occasion que, même si on ne sait pas pourquoi, vous avez un risque très élevé de décéder dans les cinq ans. Quelles seraient les implications médicales et psychologiques d'un tel résultat ? Quelles conséquences cela pourrait-il avoir pour vous sur le plan professionnel ou financier, pour l'obtention d'une assurance santé ou d'un prêt immobilier ? Au-delà de ce que l'intelligence artificielle est capable de faire, il est donc urgent de mener une véritable réflexion éthique sur son utilisation. Il ne s'agit évidemment pas d'en freiner le développement, mais de favoriser un développement bienveillant de l'IA en médecine. Nous reviendrons sur ces aspects dans le chapitre consacré aux risques et dérives.

Dans le domaine de l'imagerie cérébrale, l'intelligence artificielle ouvre de très nombreuses perspectives notamment quant à la détection automatisée des accidents vasculaires cérébraux. Une petite révolution s'est récemment produite aux États-Unis lorsque la start-up Viz.ai s'est vu attribuer un remboursement de 1 040 dollars par patient analysé avec leur solution. L'obtention de ce remboursement est une grande première dans le domaine et ouvre les portes d'une nouvelle ère en radiologie diagnostique automatisée. En effet, la plupart des modèles créés par des start-up dans le domaine de l'intelligence artificielle appliquée à la radiologie n'ont pas obtenu un tel remboursement et se sont heurtés aux contraintes de la régulation, y compris aux États-Unis. Ce que l'on appelle le *business model* de ces start-up n'était pas très bien défini. Fallait-il qu'elles vendent leurs logiciels couplés avec des machines de radiologie ou de façon indépendante ? Quelle raison pousserait les cabinets de radiologie ou les hôpitaux à s'équiper ? La start-up Viz.ai a résolu ce problème. Elle a en effet créé un système, *ContaCT*, utilisant le *Deep Learning*, capable de détecter sur un scanner une thrombose dans les vaisseaux cérébraux. Si une anomalie est détectée, elle contacte directement le spécialiste capable de traiter le problème sans l'intervention d'un radiologue. Viz.ai a pu montrer qu'en accélérant le parcours de soins de ce type de patients, qui ont besoin d'un traitement très rapidement, elle pouvait améliorer significativement la récupération d'AVC et diminuer les séquelles. Ainsi, contrairement à la majorité des autres études similaires, qui se contentaient de

détecter l'AVC, l'intelligence artificielle de Viz.ai a prouvé sa véritable utilité clinique. C'est la raison pour laquelle *Medicare*, l'équivalent de la Sécurité sociale aux États-Unis, rembourse à un hôpital une certaine somme pour chaque patient analysé par *ContaCT*. Viz.ai propose un « abonnement » annuel de 25 000 dollars en échange de son utilisation, ce qui signifie qu'un hôpital américain traitant un grand nombre de malades pour suspicion d'AVC thrombotique sera financièrement incité à utiliser *ContaCT*. Il paiera un abonnement très peu cher tout en bénéficiant de nombreux remboursements. En France, ce type de remboursement n'existe pas pour le moment, mais, si une solution utilisant de l'IA faisait la preuve de son avantage en termes d'efficacité thérapeutique, elle pourrait tout à fait relever d'un remboursement spécifique dans le cadre de la tarification à l'acte. Cela nécessiterait de mener une étude prospective puis de demander un remboursement spécifique aux autorités.

La cartographie du connectome (l'ensemble des connexions neuronales du cerveau) est une autre application intéressante du *Machine Learning* en image cérébrale. Les techniques d'imagerie fonctionnelle permettent notamment de mieux comprendre comment sont connectés les neurones. On estime que 1 000 neurones sont interconnectés par 10^{14} liaisons. Les techniques conventionnelles d'imagerie ne suffisent pas à rendre compte d'un si grand nombre de possibilités. Or certaines structures du connectome sont corrélées à des pathologies neurologiques. Il est donc important de caractériser ces connexions et ces structures. Différentes études ont utilisé l'apprentissage machine pour identifier les caractéristiques du connectome susceptibles d'expliquer une maladie d'Alzheimer, une épilepsie, un autisme, ou un syndrome d'hyperactivité. L'une des plus larges bases de données utilisées à des fins de recherche dans ce domaine est l'*Autism Brain Imaging Data Exchange* (ABIDE), qui comprend plus de mille IRM fonctionnelles. L'ABIDE est une initiative de recherche qui rassemble dix-sept laboratoires internationaux et dont le but est de mieux comprendre les mécanismes neurologiques à l'origine de l'autisme. Des réseaux neuronaux artificiels, eux-mêmes inspirés de nos cerveaux, servent aujourd'hui à comprendre notre encéphale. Quelle ironie !

L'IA est en passe de révolutionner une autre discipline médicale reposant sur la perception humaine : £ind£Anatomopathologie£/ind £l'anatomopathologie. Cette spécialité méconnue du grand public est primordiale dans la démarche diagnostique de nombreuses maladies. Un diagnostic de cancer ne peut par exemple être porté que sur un examen anatomopathologique d'une biopsie ou d'une pièce opératoire. Concrètement, cela consiste à observer au microscope les caractéristiques de tissus ou de cellules. Et ainsi poser le diagnostic ou vérifier que l'on a bien enlevé une tumeur dans sa totalité, en examinant les bords de l'exérèse chirurgicale. Après préparation, les biopsies ou la pièce anatomique peuvent être incluses

dans des blocs de paraffine et coupées en très fines tranches de quelques micromètres, puis analysées sous microscope par un médecin anatomopathologiste. Une fois de plus, comme en radiologie, le fondement de cette discipline est la capacité de percevoir et caractériser une anomalie. L'intelligence artificielle y démontre donc toute son utilité.

Son essor a notamment été rendu possible par la dissémination très large des techniques de numérisation des lames. Alors qu'auparavant la majorité des lames étaient directement observées sous microscope, elles sont dorénavant numérisées à très haute résolution pour être analysées par ordinateur. Y entraîner directement des algorithmes de *Deep Learning* devient donc tout à fait possible. Dans la majorité des cas, le recours aux réseaux neuronaux convolutifs ou aux réseaux récursifs est privilégié. Ainsi des études ont été menées pour déterminer le score de Gleason, c'est-à-dire l'agressivité d'un cancer de la prostate, de façon automatique, alors que traditionnellement l'anatomopathologiste doit classifier manuellement l'architecture des cellules cancéreuses. Le réseau sera programmé pour analyser de petites parties de chaque image et en tirer une conclusion globale[13].

Mais, au-delà de la classification, le *Deep Learning* permet la détection automatique des métastases. L'exemple d'application le plus intéressant est la détection de métastases ganglionnaires dans le cancer du sein. Dans la majorité des cas, lorsque l'on traite chirurgicalement un cancer du sein, on prélève en même temps un « ganglion sentinelle » c'est-à-dire le premier relais ganglionnaire – situé souvent au creux de l'aisselle – où des cellules cancéreuses risquent d'avoir métastasé. Le ganglion est ensuite envoyé par le chirurgien pour une analyse extemporanée, c'est-à-dire réalisée pendant l'opération chirurgicale. Si la présence de cellules cancéreuses est avérée, on peut alors compléter le geste chirurgical par un curage axillaire (ablation des autres ganglions de l'aisselle) au cours de la même anesthésie, sans réveiller la patiente et sans reprogrammer une intervention ultérieure. Afin d'améliorer la rapidité de l'analyse et les taux de détection, des équipes ont utilisé des algorithmes d'intelligence artificielle afin de détecter de façon automatique les métastases ganglionnaires. L'IA est non seulement capable de détecter les métastases mais elle montre également à l'anatomopathologiste la zone où elle pense les avoir détectées, afin de guider son regard. Bien que ce type d'approche relève pour le moment de la recherche pure, il n'est pas impossible que, dans les cinq à dix années à venir, ce type de méthode se généralise.

Les autres domaines d'application de l'intelligence artificielle en anatomopathologie sont vastes, du contrôle qualité à la normalisation des colorations des lames en passant par la génération d'images de très haute résolution. Certaines start-up, notamment américaines, telles que Page.ai, une *spin-off* du Memorial Sloan Kettering Cancer Center de New York, explorent ainsi les possibilités d'assistance et d'automatisation dans le domaine. On peut imaginer, certes de façon un peu provocante, que l'analyse des lames soit

à terme exécutée dans de grands centres automatisés et que les institutions et les hôpitaux paient des forfaits d'abonnement à ces centres pour récolter, préparer et analyser les tissus. Même si dans un premier temps tout diagnostic devra sans doute être validé par un humain, il est difficile d'estimer l'évolution et le degré d'automatisation final qui sera atteint. Cela pose bien évidemment des questions sur la fiabilité du diagnostic, mais aussi en termes de formation des futurs anatomopathologistes, car si tout est automatisé, qui possédera encore l'expertise nécessaire pour enseigner cette discipline et, *in fine*, pourra vérifier que la machine ne se trompe pas ?

L'autre problème tient au diagnostic de maladies et de tumeurs rares puisqu'il est par définition très complexe d'entraîner une intelligence artificielle efficace à partir d'un très faible nombre de cas.

En dermatologie, les réseaux neuronaux profonds ont été utilisés dans une étude publiée en 2017 dans *Nature* par une équipe de l'université de Stanford[14]. Les chercheurs ont ainsi repris une base de données d'un peu plus de 129 000 images représentant environ 2 000 maladies cutanées différentes. Le réseau neuronal *GoogleNet Inception V3*, du même type que ceux utilisés pour la reconnaissance d'objets sur des photos, avait été pré-entraîné sur 1,28 million d'images. L'équipe a ensuite comparé les performances du *Deep Learning* à celles de vingt et un dermatologues et montré que celles-ci étaient au moins aussi bonnes, voire meilleures, que celles des humains.

Lorsque Brett Kuprel, second auteur de l'article original et jeune thésard à Stanford qui a étonnamment depuis peu disparu de la circulation, a présenté les résultats de cette étude au *TensorFlow Developer Summit*, la grand-messe du *Machine Learning* de Google en 2017, je l'ai senti presque gêné d'être parvenu à publier dans *Nature* une étude aussi « simple » sur le plan technique. En effet, les chercheurs n'avaient pas eu besoin de créer un nouveau réseau neuronal mais avaient réutilisé un réseau existant, partiellement réentraîné sur une nouvelle base de données. En appliquant ces techniques, cette équipe avait obtenu des performances diagnostiques aussi bonnes, voire meilleures que les plus grands experts en dermatologie. C'est pour moi une parfaite illustration de la beauté de ce type d'approche : j'ai vraiment saisi la puissance de l'IA appliquée à la médecine en découvrant que cet étudiant en thèse d'informatique était parvenu, grâce à son ingéniosité, à résoudre un problème médical pour lequel il n'avait aucune expertise. Au cours de cette même présentation, Brett Kuprel a aussi montré un prototype de leur intelligence artificielle déployée sur son smartphone : il a utilisé sur scène l'appareil photo de son téléphone pour scanner des lésions cutanées en direct. À peine la lésion était-elle scannée que l'intelligence artificielle a proposé un diagnostic avec un pourcentage de certitude très élevé. Rendant le dépistage de lésion cutanée totalement ubiquitaire, cette application permet de répondre à l'une des grandes problématiques médicales actuelles, celle de la pénurie de médecins. Ainsi, si vous avez le moindre doute sur un grain de beauté ou un

nouveau bouton, plus besoin d'attendre six mois avant le rendez-vous chez votre dermatologue, prenez votre téléphone, scannez-vous la peau et obtenez une réponse fiable et instantanée ! Pour autant, l'intelligence artificielle ne remplacera pas le dermatologue : dans le cas de cette étude, il s'avéra que l'IA avait un très grand point faible. Elle était totalement incapable de diagnostiquer quoi que ce soit sur des peaux noires. En effet, puisqu'elle avait été entraînée sur une base de données créée à partir de peaux blanches en très grande majorité, elle n'avait aucune idée de comment reconnaître une lésion sur des peaux différentes. Cette question des biais liés aux données d'entraînement est primordiale en intelligence artificielle, comme nous le verrons plus tard. Elle peut être à l'origine de graves inégalités, voire d'erreurs médicales.

En cardiologie, l'IA joue déjà un rôle prépondérant dans l'analyse automatisée des électrocardiogrammes (ECG), un examen fondamental. L'électrocardiogramme permet de détecter et de diagnostiquer un très grand nombre de pathologies, des arythmies cardiaques aux anomalies de conduction, en passant par la fibrillation auriculaire. Il consiste à enregistrer l'activité électrique du cœur en plaçant des électrodes sur différentes zones de la cage thoracique et au niveau des extrémités des quatre membres. Pour comprendre ce dont on parle, il faut savoir qu'une partie du cœur est constituée de circuits électriques qui transmettent des impulsions au muscle cardiaque, selon une fréquence et un rythme définis en fonction des besoins en oxygène du corps.

Plusieurs entreprises proposent déjà des services d'analyse automatisée des électrocardiogrammes utilisant l'intelligence artificielle. Cardiologs, une start-up française, est l'une des pionnières. Fondée en 2014, elle a obtenu le marquage CE en 2016 et l'autorisation de la FDA en 2017. Elle commercialise depuis 2018 une solution d'IA qui permet de diagnostiquer à distance de très nombreuses pathologies cardiaques. Le taux de faux positifs dans le diagnostic de fibrillation auriculaire est divisé par dix avec une très grande sensibilité de 97 %[15]. Et l'analyse des Holters ECG, enregistrements de l'activité cardiaque généralement réalisés sur vingt-quatre heures très fastidieux à interpréter, se trouve nettement accélérée.

Au-delà de l'analyse de ces examens médicaux, l'intelligence artificielle pourrait occasionner une véritable révolution lorsqu'elle est utilisée en complément d'appareils grand public, tels les *smartwatchs*. Depuis quelques années, ces appareils peuvent enregistrer l'électrocardiogramme de leur utilisateur vingt-quatre heures sur vingt-quatre et sept jours sur sept. Étant donné que ces appareils sont très largement déployés, on peut imaginer dépister des troubles du rythme que l'on ne diagnostiquait pas avant. Cet argument (de vente) est notamment mis en avant par certaines marques, comme Apple, et on lit donc régulièrement l'histoire d'un utilisateur sauvé par sa montre qui lui aurait diagnostiqué un trouble du rythme avant que celui-ci

ne lui soit fatal. Sur le papier, cette idée est très séduisante. En réalité, les implications, notamment en termes de détections d'anomalie du rythme qui n'auront jamais de conséquences cliniques, ne sont pas encore bien comprises. Que faire en effet d'une anomalie mise en évidence par une montre ? Faut-il l'explorer, faut-il la traiter ? Avec tous les risques iatrogéniques (effets secondaires du traitement) que cela peut représenter ? Quel sens cela aurait-il de traiter une anomalie qui n'occasionne aucun symptôme, surtout si son traitement entraîne des effets secondaires ? Cela représenterait de plus un coût financier significatif. Il n'est pas du tout évident à l'heure actuelle que le dépistage de ces anomalies à grande échelle puisse réellement sauver des vies. Il pourrait même être inutilement anxiogène.

L'IA trouve également une application en gastro-entérologie pour détecter des maladies du tube digestif, qu'elles soient cancéreuses ou non. Actuellement, le dépistage et le diagnostic de ces maladies sont réalisés soit par vidéocapsule, soit par fibroscopie digestive haute ou basse (coloscopie). Le système de vidéocapsule consiste à avaler une petite capsule comprenant une caméra dont les images seront analysées une fois tout le tube digestif parcouru. Ce type de système est commercialisé depuis les années 2000 et particulièrement intéressant afin de réaliser le diagnostic non invasif de saignement digestif, des maladies inflammatoires chroniques de l'intestin ou des tumeurs. L'IA servirait ici à réduire le temps d'interprétation des images voire à automatiser la détection de lésion suspecte. Alors que l'étude d'une vidéocapsule prend en général de 50 à 60 minutes, l'intelligence artificielle pourrait réduire ce temps à une dizaine de minutes. Dans les études publiées, ce type d'approche sélectionne automatiquement les images les plus évocatrices ou les plus suspectes avant de les présenter au clinicien pour analyse. Une équipe de l'université de Tokyo a même utilisé un réseau neuronal afin de réduire ce temps à cinq minutes[16]. Le taux de détection des lésions était similaire à une lecture humaine classique. Une autre équipe de l'université de Californie Irvine a également entraîné un réseau de *Deep Learning* à partir de 53 000 images de vidéocapsule afin d'identifier des anomalies. Leur algorithme était capable d'identifier des lésions au niveau de l'œsophage, de l'estomac, de l'intestin grêle et du côlon avec une grande sensibilité. Comme dans les domaines de l'imagerie et de l'anatomopathologie, l'intelligence artificielle est donc parfaitement adaptée pour réduire le temps de lecture d'une vidéocapsule, mais aussi mieux dépister d'éventuelles lésions.

Des approches similaires ont été tentées afin de mettre l'intelligence artificielle au service de la réalisation et de l'interprétation des fibroscopies digestives hautes et basses, notamment pour déceler d'éventuels cancers. L'interprétation ne peut être donnée *a posteriori* ou même avec quelques minutes de retard car on doit interpréter en temps réel ce que l'on voit sur l'écran vidéo durant l'examen. L'IA doit être capable d'analyser en temps

réel, pendant l'endoscopie, les images qui lui sont montrées. Ce signal vidéo est par nature plus complexe à analyser que des images fixes. De nombreuses équipes ont évalué les performances d'un réseau neuronal convolutif (CNN) afin de détecter automatiquement des lésions cancéreuses ou pré-cancéreuses. Même si l'examen est toujours opérateur-dépendant, puisque le fibroscope reste manipulé par le médecin, les vidéos produites peuvent être automatiquement analysées par *Deep Learning*. Au-delà de la détection simple, l'intelligence artificielle permet de décrire automatiquement la taille et la morphologie d'un polype, voire d'évaluer les caractéristiques histologiques macroscopiques laissant supposer un cancer.

En urologie, plusieurs équipes ont entraîné un CNN à analyser en temps réel les images d'une cystoscopie pour détecter d'éventuelles petites tumeurs de la paroi vésicale qui seraient passées inaperçues. L'intelligence artificielle sert aussi à repérer des cellules de cancer de la vessie dans les cytologies urinaires ou à l'analyse d'image échographique, afin, par exemple, de mesurer automatiquement la taille de la vessie.

Le recours à l'IA est également envisagé en psychiatrie. La première application concerne l'utilisation d'agents conversationnels, également appelés *chatbots* (contraction anglophone de *chat* et *robot*), afin d'évaluer la santé mentale d'un patient, lui apporter une information médicale, voire lui proposer une psychothérapie. Le premier exemple de *chatbot* a été mis au point par Joseph Weizenbaum en 1966. Le programme, baptisé *ELIZA*, imitait un psychothérapeute et reformulait les phrases des patients discutant avec lui[17]. Il a durablement influencé le développement de ce type de dispositifs. Bien plus tard, en 1991, un second agent conversationnel psychothérapeute, *Dr Sbaitso*, est sorti sur MS-DOS. Les exemples de *chatbots* se sont multipliés depuis et, parmi ceux qui ont été évalués de façon rigoureuse, l'on peut citer *Woebot* (2017), *Koko* (2018) et *Shim* (2018). En France, la start-up WeFight crée des *chatbots* destinés à aider les patients atteints de maladies chroniques. Ces IA peuvent converser naturellement avec les patients, répondre à leurs questions et leur fournir des informations médicales. Elles leur permettent aussi d'évaluer et de suivre l'évolution de leurs symptômes, leur rappelle leur rendez-vous et prise médicamenteuse. Afin d'estimer la qualité d'une des IA créées par WeFight, nous avons mené un projet de recherche évaluant la qualité de l'information médicale fournie aux patientes par *Vik*, l'IA dédiée aux patientes atteintes d'un cancer du sein. Nous cherchions à montrer que les réponses apportées par un *chatbot* aux questions des patientes étaient au moins aussi satisfaisantes que celles données par un comité de médecins pluridisciplinaire (chirurgien oncologue, oncologue médical et oncologue radiothérapeute). Un pool des vingt questions les plus fréquemment posées a été préalablement défini à partir des patientes utilisant l'application de WeFight. Ces questions ont été posées d'une part au *chatbot*

et d'autre part à un comité médical pluridisciplinaire. Un groupe de 140 patientes a ensuite été divisé en deux : une partie a reçu les réponses du *chatbot*, tandis que l'autre recevait les réponses du comité de médecins. Les patientes ne savaient pas si elles recevaient les réponses des médecins ou du *chatbot*. Elles ont enfin répondu individuellement à un questionnaire standardisé afin d'évaluer la qualité de l'information médicale. Le taux de satisfaction a été de 69 % dans le groupe *chatbot* contre 64 % dans le groupe médecins. Il s'agit de la première étude à avoir directement comparé l'information délivrée par un *chatbot* à celle donnée par des médecins de façon prospective, randomisée et en aveugle[18].

Toujours en psychiatrie, l'IA pourrait servir à la détection et au diagnostic automatisé de syndrome dépressif et du risque suicidaire, en utilisant par exemple les données des réseaux sociaux. L'une des études les plus importantes a été publiée par une équipe de l'université de Princeton en 2018 dans le prestigieux journal *PNAS* (*Proceedings of the National Academy of Sciences of the United States of America*). Les chercheurs ont entraîné une intelligence artificielle à partir de l'historique des statuts Facebook de 683 patients, dont 114 patients dépressifs, en utilisant des techniques de traitement du langage naturel, un sous-domaine du *Machine Learning*. Ils ont ainsi pu dépister la majorité de leurs patients effectivement dépressifs jusqu'à trois mois avant que les premiers symptômes ne soient documentés sur le plan médical[19]. Plus original encore, une équipe du département de psychologie d'Harvard a utilisé 43 950 photos Instagram de 166 personnes afin d'identifier les personnes les plus à risque de dépression. Les chercheurs ont montré que certains filtres et certaines couleurs des photos étaient fortement prédictifs de dépression[20]. Ce type d'applications est particulièrement intéressant car il touche à la problématique de la balance bénéfice-risque de l'utilisation de l'IA en santé. En effet, être capable de dépister des syndromes dépressifs, voire des idées suicidaires, permettrait éventuellement d'intervenir avant un geste fatal et donc de sauver des vies. Cependant, ce type de solutions implique une surveillance globale qui rappelle évidemment les pires dystopies, comme celle décrite dans le roman de George Orwell, *1984*. Concrètement, faut-il protéger les personnes suicidaires en sacrifiant nos vies privées ?

Un nombre croissant de dispositifs médicaux (DM) exploitant l'IA (DM-IA) sont commercialisés chaque année. En Europe, le nombre de DM-IA sur le marché était de 12 en 2018 et de 100 en 2019. Aux États-Unis ce chiffre est passé de 9 à 80 sur la même période. La façon dont ces DM-IA sont approuvés en Europe et aux États-Unis n'est actuellement pas très claire, alors que la commercialisation des DM sans IA fait l'objet d'une régulation passant par l'obtention d'un marquage CE en Europe ou d'une approbation par la FDA aux États-Unis. Bien qu'un DM-IA défectueux puisse présenter un risque réel pour les patients, les méthodologies d'évaluation seraient potentiellement moins rigoureuses que pour d'autres catégories de DM.

Une équipe de l'université de Zurich a récemment publié une étude sur les DM-IA approuvés depuis 2015 en Europe et aux États-Unis dans le *Lancet Digital Health*[21]. En fouillant les bases de données des agences de régulation européennes et américaines, ils ont respectivement identifié 240 DM-IA approuvés en Europe et 222 aux États-Unis. La majorité de ces dispositifs est utilisée en radiologie, en cardiologie et en neurologie. La plupart sont destinés aux professionnels de santé et environ 20 % peuvent être directement utilisés par les patients. La très grande majorité de ces dispositifs est vendue par de petites entreprises, tandis que 20 % sont distribués par des entreprises plus importantes déjà bien implantées dans le domaine de la santé. Certains dispositifs sont approuvés aux États-Unis mais pas en Europe et vice-versa, ce qui pose la question des critères précis d'approbation ou non d'un DM-IA.

En Europe, contrairement aux États-Unis, les dispositifs médicaux sont classés en fonction de leur risque potentiel pour la santé sur une échelle de 1 (risque le plus faible) à 4 (risque le plus élevé). Des 240 DM-IA approuvés en Europe, 35 % étaient en classe 1, 50 % en classe 2 et 1 % en classe 3. Aucun en classe 4. Or, la réalisation d'essais cliniques n'est nécessaire que pour la régulation des DM-IA de classe 4. Nous prenons donc le risque que des dispositifs utilisés pour certaines tâches très importantes, comme diagnostiquer des cancers du sein ou des hémorragies cérébrales, n'aient pas été évalués rigoureusement et présentent un danger pour les patients.

Une étude, publiée récemment dans le journal *Nature Medicine* par une équipe de l'université de Stanford[22], révèle que, sur 130 DM-IA, presque aucun n'a été évalué. Pis, aucun des 54 DM-IA considérés à haut risque n'a été apprécié dans une étude prospective. Ces chercheurs ont également montré que la grande majorité de ces dispositifs n'a été testée que sur une seule population, au risque de biaiser l'algorithme en raison du surapprentissage (l'IA n'a appris que d'une population et est incapable de fonctionner sur une autre) et d'exposer le patient à une erreur.

Les DM-IA doivent faire l'objet d'évaluations prospectives sur plusieurs hôpitaux, afin de vérifier la fiabilité de l'algorithme sur différentes populations. Il faudra de plus exercer une surveillance de tous les dispositifs après leur autorisation de mise sur le marché (AMM), afin d'écarter les risques de biais et de mauvaises performances non détectés dans les études prospectives initiales, à l'image des processus de pharmacovigilance que l'on utilise pour les médicaments ou des phases 4 menées après l'AMM d'un nouveau médicament.

4

Traiter grâce à l'IA

La chirurgie sera sans doute l'une des dernières spécialités à être automatisée. Cela tient bien évidemment à ce qu'il s'agit d'une spécialité très manuelle et dépendante du chirurgien. Cependant, l'avènement de la chirurgie robotique, permettant au chirurgien de piloter des instruments à distance avec une vue 3D du site opératoire, ouvre la brèche et permet d'envisager à moyen terme peut-être pas une automatisation mais au moins une assistance par des solutions d'intelligence artificielle.

Les toutes premières étapes nécessaires à l'automatisation ont déjà été franchies, notamment dans le cadre de travaux publiés par l'équipe de Fei-Fei Li à Stanford. Cette équipe a travaillé sur la reconnaissance automatisée des différents outils chirurgicaux sur des images de laparoscopie[23]. Ils ont pour cela traité le signal vidéo image par image grâce à un réseau neuronal et montré qu'ils étaient capables d'identifier correctement les pinces, ciseaux et autres agrafeuses utilisés en chirurgie. Les chercheurs ont même entraîné une IA à évaluer la qualité du geste chirurgical grâce à une échelle standardisée appelée GOALS (*Global Operative Assessment of Laparoscopic Skills*) et montré que l'IA évaluait le geste chirurgical aussi bien qu'un groupe de chirurgiens experts.

Ce travail pionnier est certes préliminaire mais permet d'envisager, au-delà de la simple reconnaissance des outils chirurgicaux, de passer à l'étape suivante qui consisterait à piloter automatiquement ces outils et enfin à réaliser une opération chirurgicale entièrement robotique. Accepteriez-vous un jour d'être opéré par une machine ?

Autre spécialité en passe de se trouver partiellement automatisée : la radiothérapie, principal traitement du cancer avec la chimiothérapie et la chirurgie. La radiothérapie est l'une des rares disciplines médicales où le traitement administré est documenté de façon très précise et complète *via* un système informatique qui enregistre en permanence les événements. Ces informations peuvent être extraites et leur pertinence analysée pour effectuer une prédiction. Dans cette discipline, qui repose beaucoup sur l'imagerie et l'informatique, l'IA servirait à automatiser chaque étape du traitement, notamment sa préparation : avant de débuter sa radiothérapie, tout patient bénéficie d'un scanner qui est ensuite importé sur le système informatique de planification. La tumeur et les organes l'entourant sont alors « contourés » manuellement par l'oncologue en trois dimensions. Ce procédé long et

fastidieux, appelé la segmentation, peut déjà être automatisé par *Deep Learning*. Les publications dans le domaine sont nombreuses pour les tumeurs du cerveau, des poumons, de l'abdomen et du pelvis. Les performances obtenues sont correctes et ce qui prenait avant plusieurs heures est accompli par l'IA en quelques secondes. L'un de mes projets de recherche à Stanford a justement consisté à entraîner une IA à « contourer » automatiquement la prostate et comparer les résultats à ceux obtenus par des médecins radiothérapeutes. Nous avons ainsi montré que l'IA était tout à fait capable d'exécuter cette tâche, y compris sur des scanners de mauvaise qualité[24]. Nous y gagnerions en temps et en précision.

Après la segmentation, une dosimétrie permet d'administrer la dose prescrite sur la tumeur, sans dépasser certains seuils sur les organes voisins (on parle de contraintes de dose). Ainsi, pour une même maladie, le traitement planifié différera significativement selon l'anatomie du patient. Cette étape de calcul consiste schématiquement à modeler la trajectoire des faisceaux d'irradiation pour la façonner parfaitement autour de la tumeur. Cette tâche s'effectue manuellement pour le moment (avec assistance de l'ordinateur), mais gagnerait à être entièrement automatisée afin d'obtenir le plan d'irradiation qui traite le mieux la tumeur, sans endommager les organes sains voisins. Plusieurs dizaines d'études publiées en démontrent la faisabilité.

Une fois le traitement planifié, les séances débutent. Là encore, tout est numérisé : chaque jour, le positionnement du patient sous l'accélérateur de particules fait l'objet d'un contrôle, qui peut occasionner des corrections dans les trois axes de l'espace (+/– les rotations) afin d'être sûr d'irradier la bonne zone. L'ensemble de ces paramètres (date, heure des séances, décalage table) est enregistré de façon automatique dans un système informatique, appelé *Record and Verify*, utilisé au quotidien. Ce système permet de contrôler et de quantifier la correspondance entre le traitement planifié et celui effectivement réalisé. Cette particularité rend l'exploitation des données en radiothérapie très intéressante. Il est ainsi possible de nourrir des algorithmes capables de prédire les effets secondaires des traitements (afin de les minimiser) ou les chances de guérison.

5

Prévenir grâce à l'IA

L'IA peut être utilisée avant même la naissance pour optimiser la médecine reproductive. Dans ce domaine, l'analyse automatisée d'images par IA a deux applications directes : l'analyse de la quantité et de la qualité des spermatozoïdes et des ovocytes et l'analyse de la qualité des embryons obtenus par fécondation *in vitro*. La qualité des spermatozoïdes permet de prédire la fertilité d'un homme. À l'heure actuelle, les principales caractéristiques telles que le nombre, la mobilité, la vitalité et la morphologie sont évaluées manuellement sous microscope par un biologiste. Là encore, certaines équipes se sont servies d'un réseau neuronal, le *VGG16*, ordinairement utilisé pour la reconnaissance d'objets sur des photos, pour prédire la qualité des spermatozoïdes avec une précision de 80 %. Les meilleurs spermatozoïdes sont ainsi sélectionnés afin d'augmenter les chances de succès de fécondation *in vitro*.

Après fécondation *in vitro*, et avant que les embryons ne soient injectés à la future mère, une nouvelle sélection doit permettre d'augmenter les chances de succès de la nidation et de poursuite de la grossesse. Étant donné qu'un peu plus de 10 % de la population mondiale rencontre des difficultés pour concevoir naturellement, ces questions autour de la fécondation *in vitro* et du rôle de l'intelligence artificielle sont centrales. Le taux de succès d'une fécondation *in vitro* étant bas, il est légitime d'utiliser l'ensemble des techniques à notre disposition pour augmenter les chances de succès et éviter des souffrances inutiles à ceux qui désirent un enfant. Une équipe de l'université de Weill Cornell aux États-Unis a ainsi entraîné un réseau neuronal profond de type *Inception V3* à reconnaître les meilleurs embryons à partir de 50 000 images. Le système de *Deep Learning* a obtenu de meilleures performances que les biologistes participant à l'étude[25].

Le *Machine Learning* a aussi un rôle à jouer dans la prévention du risque de diabète gestationnel, diagnostiqué la plupart du temps entre la 24e et la 28e semaine de grossesse. Une détection plus précoce pourrait considérablement réduire le risque de complications. Plusieurs études l'ont démontré. En 2020, une équipe du Weizmann Institute of Science en Israël a ainsi entraîné un modèle de prédiction à partir des dossiers de plus de 580 000 patientes. Les résultats, excellents, ont permis d'anticiper 80 à 85 % des diabètes gestationnels à partir de neuf questions simples. Comme dans d'autres travaux précédemment cités, les auteurs ont utilisé les valeurs de Shapley afin de fournir les raisons pour lesquelles l'IA avait donné telle ou telle prédiction

pour chaque patiente[26].

En pédiatrie, l'intelligence artificielle a principalement été utilisée afin de prédire la prématurité. Selon la plupart des études, le taux de naissance prématurée dépasse actuellement 10 %. Étant donné que cette prématurité est à l'origine d'une morbidité et d'une mortalité significatives, le dépistage et le diagnostic précoce de ces maladies sont un enjeu de santé publique. Les premiers systèmes algorithmiques développés afin de prédire le risque de prématurité remontent aux années 1980 avec les travaux de Creasy, puis de Woolery en 1994. Plus récemment, des équipes ont utilisé les images d'échographie du second semestre afin de corréler le périmètre crânien et la longueur fémorale au développement des fonctions cognitives.

Autre domaine où l'IA permet de prévenir des complications, la néphrologie, spécialité qui s'intéresse à la fonction rénale : une étude emblématique a été publiée en 2019 dans *Nature* par une équipe de DeepMind, la start-up londonienne dont nous avons déjà parlé, rachetée en 2014 par Google. Les chercheurs ont développé une méthode fondée sur le *Deep Learning* afin de prédire, de façon continue dans le temps, une éventuelle détérioration de la fonction rénale parmi une cohorte de patients hospitalisés. Le modèle, nourri des dossiers médicaux de plus de 700 000 patients, extraits de plus de 2 200 centres de soins, a permis de prédire 90 % des insuffisances rénales aiguës nécessitant une dialyse, le tout quarante-huit heures à l'avance[27]. Une approche très intéressante si l'on considère qu'environ 11 % des décès à l'hôpital sont liés à l'absence de détection et d'anticipation d'une détérioration de l'état clinique d'un patient par son médecin. Bien que cette étude ait fait grand bruit à l'époque de sa publication, elle illustre également un paradoxe propre à l'intelligence artificielle appliquée à la médecine. En effet, être capable de prédire deux jours à l'avance l'apparition d'une insuffisance rénale reste une première étape, insuffisante en tant que telle, puisqu'elle n'implique pas d'être capable d'empêcher cette détérioration. Concrètement, cela ne démontre pas une amélioration des chances de guérison ou de survie du patient. Il est donc primordial, au-delà de la création de systèmes algorithmiques techniquement sophistiqués, de chercher à résoudre de véritables problèmes cliniques et de démontrer que l'intelligence artificielle permet effectivement de les améliorer. Pour ce faire, le développement d'essais cliniques prospectifs dédiés sera nécessaire, comme ceux menés avant toute commercialisation d'un nouveau médicament. Nous développerons cette thématique dans le chapitre consacré à la régulation de l'IA en santé.

6

Le citoyen augmenté

On estime que les maladies chroniques non infectieuses constituent la principale cause de morbidité et de mortalité dans le monde. Ces maladies résultent la plupart du temps de comportements dangereux ou de consommation excessive. L'adoption d'un mode de vie plus sain pourrait donc prévenir leur apparition. Ici, l'intelligence artificielle a aussi un rôle à jouer : en aidant à corriger certains comportements ou en alertant lorsqu'une consommation excessive d'aliments ou de substances menace la santé d'une personne.

L'application la plus directe concerne bien évidemment les régimes alimentaires. Un régime alimentaire déséquilibré favorise le risque d'obésité, de diabète, de maladies métaboliques et augmente la mortalité notamment cardiovasculaire. Des recommandations diététiques sont à même de réduire significativement le risque de décéder à cause de ces maladies. Cependant, de nombreuses personnes éprouvent des difficultés importantes à les suivre. Une équipe italienne a donc mis au point une IA capable de compenser de façon dynamique certains comportements par des comportements sains[28]. Par exemple, si un utilisateur mange du gâteau le lundi, l'IA lui propose de manger plutôt des légumes le lendemain. L'IA n'informe pas simplement sur ce qui est sain ou pas (ça, tout le monde le sait !) mais aide l'utilisateur à s'y tenir. Cette flexibilité et cette adaptabilité permettraient théoriquement de favoriser l'adhésion à un régime alimentaire équilibré. Au lieu de culpabiliser l'utilisateur, le système aide à compenser les apports caloriques afin d'équilibrer l'ensemble sur une période donnée, de façon personnalisée. D'autres travaux vont encore plus loin dans la prédiction notamment de la satiété et des niveaux de glycémie après un repas. Une équipe menée par David Zeevi, de l'université Rockefeller, a ainsi mis au point une IA capable de fournir ses prédictions à partir des données de huit cents participants qui devaient renseigner leurs habitudes alimentaires, leur activité physique, leurs mensurations, et une analyse de leur microbiote intestinal[29]. Évidemment, entrer l'ensemble de ces informations risque de se révéler rapidement difficile. Toutes les personnes qui ont utilisé des applications destinées à compter les calories le savent. C'est la raison pour laquelle McAllister et son équipe ont développé un réseau neuronal profond qui collecte ces informations, non plus en entrant manuellement le type et la quantité de nourriture ingérée, mais en permettant à l'utilisateur de prendre une photo de son assiette sans aucune autre information. Des applications mobiles ont depuis repris ce principe

(*LifeSum*, *Snap It*, *Calorie Mama*, *BiteSnap*).

Le deuxième domaine dans lequel l'intelligence artificielle aide à réduire les comportements néfastes est celui de l'activité physique et du fitness. Si les applications que nous venons de citer permettent d'enregistrer vos apports caloriques, il est facile de comprendre qu'il est tout aussi important d'enregistrer vos dépenses énergétiques, afin d'être à l'équilibre.

De nombreux dispositifs sont déjà disponibles, les plus utilisés étant les smartphones et les *smartwatchs*. Si ce type d'appareil mesure aujourd'hui le nombre de pas que l'on fait, il faut savoir que c'est grâce à un modèle d'intelligence artificielle entraîné pour corréler les mouvements enregistrés sur les accéléromètres et la dépense énergétique ou le nombre de pas effectués dans une journée. Des agents conversationnels ont également été développés afin d'encourager leurs utilisateurs à avoir une activité physique régulière. Ces agents peuvent être des applications de type *chatbot* ou des agents vocaux sur enceintes connectées[30].

L'IA joue aussi un rôle dans l'évaluation et l'amélioration de la qualité et de la quantité du sommeil. La qualité du sommeil, essentielle pour la santé, est en général évaluée à l'aide d'une polysomnographie qui peut être corrélée au taux d'oxygène sanguin et à la respiration. Comme pour le suivi de l'activité physique, les smartphones et surtout les *smartwatchs* sont beaucoup utilisés. Certains algorithmes permettent ainsi de reconnaître l'apnée du sommeil ou un sommeil agité.

Concernant la prise en charge de la fatigue, l'université de Cornell et Google ont collaboré à une étude explorant l'intérêt d'un algorithme de *Machine Learning* appliqué au *tracking* du regard sur smartphone. Les utilisateurs devaient accomplir pendant une heure une série de tâches face à un téléphone, la direction de leur regard étant enregistrée. Les chercheurs ont ensuite comparé les motifs des mouvements oculaires à des questionnaires de fatigue validés et montré une très forte corrélation entre les deux[31]. Il est donc facile d'imaginer qu'à terme ce type d'algorithmes soit déployé sur votre téléphone ou dans votre voiture, afin de détecter les moments de fatigue et vous encourager à vous reposer.

III

Perspectives

1

De la Pascaline au smartphone

Début février 2020, il ne me reste plus que trois mois pour terminer mes recherches à Stanford. Afin de remettre les choses en perspective et prendre un petit peu de recul, nous décidons de visiter en famille le Computer History Museum (CHM), à deux pas du campus de Google à Mountain View. Le CHM retrace l'histoire de l'informatique, de la Pascaline (la machine à calculer de Blaise Pascal) à l'iPad d'Apple.

Au fil des vitrines et des couloirs, je mesure le chemin parcouru dans ce domaine et à quel point l'informatique nous aide désormais à exécuter de très nombreuses tâches. J'ai aussi pleinement conscience de notre dépendance vis-à-vis de la technologie, toujours plus performante. Si demain tous les ordinateurs tombaient subitement en panne, nous serions littéralement paralysés et ne pourrions remettre le monde en ordre de fonctionnement avant plusieurs années. Ainsi, même si je sais programmer des algorithmes afin de déterminer un diagnostic ou le pronostic d'une maladie, je suis totalement incapable de recréer ou de réparer le processeur d'un ordinateur. Cette perspective me terrifie et je m'interroge sur les moyens préventifs à mettre en place afin d'empêcher ou de pallier cette dépendance extrême. Au fil des époques retracées dans le musée, nous retrouvons les calculateurs utilisés dans le domaine militaire pour piloter les systèmes de guidage de missiles dans les années 1960 et 1970, les super-ordinateurs Cray et certains appareils exotiques tels que les robots policiers du Sunnyvale Police Department. Petit à petit, nous parvenons à une époque que je connais : les années 1980. Je croise ainsi ce par quoi tout a commencé pour moi : la console Nintendo Entertainment System et l'Amstrad CPC 6128 ! Le musée possède également des Apple I et II. Il est d'ailleurs assez incroyable de voir un exemplaire original de l'Apple II sachant que les rares pièces encore disponibles valent plusieurs centaines de milliers de dollars. Je mesure le travail accompli à la fin des années 1970 par Steve Wozniak et Steve Jobs, les fondateurs d'Apple. Le chemin parcouru depuis est vertigineux. Je continue d'avancer dans les allées, je regarde ma fille Anaïs, née en 2017, qui n'aura jamais connu le monde d'avant l'informatique et Internet. Il me semble alors évident que nous n'avons pas totalement conscience des changements civilisationnels que nous vivons. Nous sommes maintenant dans la salle consacrée aux années 2000 : après 2007 et la commercialisation du premier iPhone, la majorité des ordinateurs exposés au CHM sont en réalité des terminaux mobiles ou des tablettes. En parallèle, le musée ne présente plus de machines mais des

entreprises et donc plutôt les logiciels qu'elles éditent. On peut par exemple voir les premiers objets *brandés* (personnalisés aux couleurs de la marque) des fondateurs et ingénieurs de Google.

En ce sens, le musée illustre parfaitement le passage du *hardware* (matériel) au *software* (logiciel), en ligne directe avec les propos tenus dès 2011 par Marc Andreessen, cofondateur de la firme d'investissement Andreessen Horowitz : « Le logiciel mange le monde » (*Software is eating the world*). La visite se termine par l'un des premiers véhicules de *Waymo*, *spin-off* de Google spécialisé dans les voitures autonomes. L'exemplaire date de 2016, c'est-à-dire quatre ans plus tôt, soit une éternité à l'échelle de la Silicon Valley. Cette voiture pouvait contenir deux personnes mais depuis le système a été adapté à un très grand van Chrysler, le Pacifica, hybride évidemment, que l'on croise tous les jours et partout dans la vallée. Avant de quitter le musée, les visiteurs peuvent laisser un message sur un tableau fixé dans le hall. Je prends un stylo et écris sous forme de clin d'œil « *import tensorflow as tf* », l'une des premières lignes de code d'un projet d'intelligence artificielle permettant d'importer le *framework Tensorflow*, bibliothèque de *Deep Learning* de Google.

Après avoir examiné les possibilités offertes par l'IA dans chaque spécialité, il est temps de discuter de son évolution prévisible : comment s'intégrera-t-elle à votre smartphone ? Comment deviendra-t-elle omniprésente et invisible ? Enfin, comment transformera-t-elle le système de soins ? Quels sont les risques potentiels liés à son utilisation, notamment pour le secret médical ? Et quels seront, selon moi, les leaders de demain dans le domaine. J'exposerai la meilleure façon de gérer les données de santé pour ne pas freiner le développement de l'IA, tout en protégeant les patients. La question centrale de la régularisation de l'IA sera aussi abordée. Je vous proposerai enfin deux scénarios futuristes, l'un à moyenne échéance (2031) et l'autre à plus long terme (2041).

2

Du creux de votre main à l'air qui vous entoure

Les applications couplées à l'IA sur mobile sont vouées à prendre une place de plus en plus importante dans la pratique médicale. Elles peuvent notamment être utilisées afin de définir la décision thérapeutique. On estime qu'environ 15 % des applications actuellement disponibles sont capables de guider les stratégies de traitement. Parmi elles, les applications utilisant les caméras de téléphone sont majoritaires. L'exemple typique est la détection de lésions cutanées telles que décrites dans la partie consacrée à la dermatologie. D'autres applications plus sophistiquées nécessitent par exemple l'ajout d'un minimicroscope au téléphone. Certaines équipes travaillent même à l'ajout d'un scanner rétinien afin de faciliter le dépistage et le diagnostic de la rétinopathie diabétique de façon automatique et décentralisée. Il faut cependant garder à l'esprit que la validation clinique de ces différentes applications est très insatisfaisante. Les utiliser hors du cadre de la recherche clinique est donc déconseillé.

Le deuxième type d'applications disponibles correspond aux algorithmes de recommandations de bonnes pratiques. Il s'agit d'intelligences créées non par apprentissage machine, mais plutôt par transcription manuelle du savoir d'un expert sous la forme de règles simples. De nombreuses études consacrées à ce type d'applications ont comparé les résultats cliniques avant et après leur utilisation et ont conclu à une amélioration. Par exemple pour optimiser l'utilisation des antibiotiques ou encore améliorer la gestion d'une hypertension ou du diabète[32].

Le domaine le plus prometteur est probablement celui de la modélisation prédictive. À partir de la récolte de symptômes simples, il est possible de prédire et de détecter la rechute d'un cancer et d'en améliorer les chances de guérison[33]. On pourrait même imaginer qu'à terme la récolte des symptômes, entrée aujourd'hui manuellement par les patients, puisse être effectuée de façon totalement passive par le smartphone, en enregistrant par exemple le poids, les mouvements et l'activité physique du patient grâce à l'accéléromètre.

À l'opposé de ce que l'on pourrait appeler l'intelligence artificielle mobile (dans votre poche), une autre forme se développe, véritablement ubiquitaire celle-ci : l'*intelligence artificielle ambiante*. Ce concept repose sur

l'utilisation conjointe de l'apprentissage machine et de capteurs intégrés à l'environnement. Déjà utilisée pour la conduite autonome de véhicules, l'IA ambiante n'a pas encore été déployée à grande échelle dans les hôpitaux. Pourtant, elle apporterait une aide précieuse au personnel soignant. Lors d'une période de pandémie par exemple, l'intérêt d'utiliser une *intelligence artificielle ambiante* dans les services de soins intensifs paraît évident. La surveillance automatique de patients hospitalisés permettrait de diagnostiquer plus systématiquement les pathologies liées à une immobilisation prolongée. Grâce par exemple à l'utilisation de caméras ou de capteurs LIDAR, les études démontrent que cette utilisation pourrait réduire de 40 % ce type de pathologies. Une étude pionnière dans le domaine a notamment placé des capteurs dans une chambre de soins intensifs afin d'enregistrer 362 heures de données chez huit patients. Un algorithme a ensuite été entraîné pour classifier les différentes activités du patient avec une grande précision. Une seconde étude plus importante, réalisée par une équipe de Stanford, a enregistré les activités de patients dans huit chambres différentes et obtenu un algorithme capable de détecter automatiquement les mouvements des patients[34]. Ce type d'approches offre la possibilité de détecter les patients les moins mobiles durant leur hospitalisation, afin de réaliser des interventions de mobilisation pour prévenir par exemple l'apparition d'escarres.

Mais la surveillance ne serait pas réservée aux patients ! Il est également envisageable, et cela a été montré par différentes équipes, de surveiller l'exécution des gestes d'hygiène par les équipes soignantes afin de prévenir l'apparition de maladies nosocomiales. Ces infections touchent chaque année environ cent millions de patients dans le monde. On estime même que jusqu'à 30 % des patients hospitalisés en soins intensifs seront touchés par une infection nosocomiale. Il s'agit donc d'un enjeu majeur de santé publique. Parmi ces études, une équipe a ainsi utilisé des capteurs au-dessus des distributeurs de solution hydroalcoolique et de savon afin de vérifier leur bonne utilisation. Ils ont ensuite comparé les performances de la détection automatique à celle d'un observateur humain et montré que les performances de la méthode automatique étaient bien meilleures (75 % *vs* 63 % de détection)[35]. Semmelweis, l'inventeur de l'hygiène, serait sans doute enthousiaste !

De la même façon, l'IA *ambiante* pourrait surveiller le déroulement d'une intervention chirurgicale. Une étude a ainsi comparé l'évaluation d'une prostatectomie effectuée par une IA ou par un groupe de chirurgiens en 2017. La concordance était excellente (92 %). Mais allons encore plus loin. L'une des procédures les plus importantes au bloc opératoire est le comptage des instruments, des compresses et autres éponges chirurgicales. Les oublier dans le corps du patient n'est pas fréquent, mais arrive et porte même un nom, voire deux : le *gossypiboma* ou textilome ! Là encore, le comptage automatisé par IA d'éponges chirurgicales à l'aide d'un code-barres ou d'une puce RFID a permis de réduire le temps nécessaire au décompte des instruments et le

nombre d'erreurs de 55 %.

Dans le domaine de la consultation médicale, l'IA *ambiante* permettrait de transcrire en temps réel le contenu d'une consultation afin de générer des comptes rendus de qualité. Une équipe de chercheurs a ainsi entraîné un réseau neuronal profond à partir de 14 000 heures d'audio de 90 000 conversations entre des patients et leur médecin. La précision de transcription du modèle était de 80 % (supérieure à la précision humaine : 76 %)[36].

L'IA *ambiante* servirait aussi à veiller sur les patients âgés en perte d'autonomie ou institutionnalisés, en détectant par exemple une éventuelle chute.

Cependant, la frontière entre bénéfice potentiel pour le patient et surveillance généralisée est très ténue. On peut citer à titre d'exemple une étude où des chercheurs ont analysé par IA les données audio et vidéo de 69 individus obtenues au cours d'entretiens de 30 minutes afin de dépister d'éventuels signes de schizophrénie. Les résultats obtenus sont une fois de plus excellents, avec une valeur prédictive positive de 95 % et une sensibilité de 84 %, mais aussi particulièrement effrayants[37].

L'un des principaux risques de l'IA *ambiante* serait évidemment son application à un domaine extra-hospitalier, notamment à la surveillance de masse dans les rues. Elle est déjà pratiquée dans des villes d'Asie du Sud-Est, comme Hong Kong. Différentes méthodes permettent d'empêcher de telles dérives mais devront être implémentées directement dans les systèmes, comme le floutage automatique des visages ou le masquage du corps.

3

La transformation du système de santé

Au-delà de ses applications concrètes dans chacune des spécialités médicales, l'IA occasionnera une révolution de l'organisation globale des soins, qu'elle se trouve dans votre poche, dans votre montre, dans votre enceinte connectée ou partout autour de nous.

Pour illustrer ce phénomène, je vous propose de prendre l'exemple de la gestion de la pandémie de Covid-19. Cette crise est particulièrement difficile pour tous : il s'agit d'être capable de freiner, et à terme de stopper la propagation du coronavirus, tout en vivant le plus normalement possible, afin d'en limiter l'impact social et économique. Pour atteindre cet objectif, une gestion efficace doit remplir cinq objectifs : la prévention et l'anticipation des contaminations, le diagnostic facile, fiable et rapide, le suivi du patient, la gestion des capacités d'accueil, des personnels soignants et des équipements et enfin, à terme, la gestion des phases de déconfinement. Dans chacun de ces domaines, des équipes de chercheurs ont mené des études destinées à évaluer des solutions d'intelligence artificielle.

En matière de prévention et d'anticipation des contaminations, l'IA a été utilisée pour cartographier la propagation du virus, notamment par des équipes de l'université Johns Hopkins. Le *Machine Learning* a aussi été utilisé pour identifier les facteurs d'évolution de la propagation du virus, analyser et exploiter la littérature scientifique (foisonnante dans le domaine), afin d'évaluer par exemple la possibilité de trouver de nouveaux traitements. L'IA a également servi à modéliser la composition et la structure en 3D du virus, à identifier des molécules efficaces contre celui-ci et à modéliser et prédire la réponse immunitaire.

Pour le diagnostic et le suivi des patients, les solutions développées sont encore plus nombreuses. L'IA aide à diagnostiquer l'infection en fonction des caractéristiques de la toux du patient, interpréter les tests de dépistage, prédire l'évolution de l'état de santé du patient à partir d'un scanner, ou encore suivre à distance les patients. Elle sert aussi à optimiser la distribution et l'analyse des tests de dépistage et de diagnostic. On ne s'en rend donc pas forcément compte, mais bien souvent, derrière les tests de diagnostic, de multiples algorithmes tournent en coulisse pour améliorer ces processus. Par ailleurs, dans les solutions de suivi à distance des patients atteints de forme non grave, telles que celle mise en place à l'Assistance publique – Hôpitaux de Paris (*Covidom*), des algorithmes comparent de façon dynamique les symptômes des patients, afin de détecter une aggravation et d'être capable de rediriger les

patients vers l'hôpital.

L'IA a également été utilisée pour lutter contre la propagation des *fake news* sur les réseaux sociaux (notamment Facebook), mesurer la capacité d'accueil en temps réel des services des urgences et des hôpitaux, prévenir la propagation du virus au sein du personnel soignant, concevoir des équipements de protection réutilisables et enfin distribuer les marchandises et matériaux essentiels.

En réalité, les algorithmes sont déjà utilisés à tous les niveaux du système de soins. On pourrait argumenter que les difficultés de gestion sont peut-être précisément liées à leur utilisation et démontrent leur inefficacité. Mais on remarquera que les pays où ces solutions ont été déployées le plus tôt possible, dès le début de la crise, dans un très grand nombre de domaines, par exemple en Corée du Sud ou à Taïwan, sont ceux où la crise sanitaire a été la plus courte.

4

Risques et dérives potentiels de l'intelligence artificielle appliquée à la santé

Puisque l'intelligence artificielle est entraînée à partir de données humaines, des biais peuvent être reproduits dans nos algorithmes si nous n'y prenons garde. Comme nous l'avons vu, n'utiliser que des images de peaux blanches pour entraîner un algorithme à reconnaître des lésions cutanées rend l'algorithme inutilisable pour des peaux noires et introduit donc par définition des inégalités.

L'utilisation d'algorithmes a déjà des effets négatifs, notamment aux États-Unis. En 2019, une équipe de l'université de Berkeley a publié une étude montrant que les algorithmes utilisés dans le système de soins américain pour définir le risque d'un patient, et donc définir les montants alloués aux soins, étaient fortement biaisés au détriment des patients afro-américains. Les auteurs de l'étude ont prouvé que, au même niveau de risque défini par l'algorithme, les patients afro-américains étaient en réalité plus malades que les patients caucasiens. Le biais réduisait de moitié le nombre de patients afro-américains à même de bénéficier de soins adaptés à leur risque[38]. La correction de ces biais est donc un enjeu majeur de l'intelligence artificielle et médicale de demain. Sans que nous nous en rendions compte, le système risque de reproduire et, avec le temps, d'amplifier les inégalités. Pour cette raison, l'explicabilité, ou ce que l'on appelle aussi l'interprétabilité, c'est-à-dire la capacité à comprendre les résultats d'un algorithme, est primordiale.

Il faut aussi envisager la possibilité que certaines personnes ou organisations utilisent l'intelligence artificielle à des fins malveillantes. Être en mesure de définir le pronostic du malade a un intérêt médical évident. Mais cette capacité risque d'être détournée pour augmenter par exemple le prix d'une assurance santé ou refuser l'assurance d'un crédit immobilier. Pour revenir à l'*intelligence artificielle ambiante*, la frontière entre un système qui monitore l'activité d'un patient afin de réduire son risque d'escarre et un système de surveillance massive est ténue. Techniquement, il s'agit exactement des mêmes systèmes. Leur seule différence tient à l'intention de l'utilisateur.

Le danger d'une utilisation discriminatoire de l'intelligence artificielle est bien réel. En 2017, une équipe de Stanford a, dans le cadre d'une étude, utilisé les caractéristiques faciales de plus de 35 000 personnes, extraites par *Deep*

Learning, pour déterminer leur orientation sexuelle. L'algorithme a été capable de classer correctement jusqu'à 91 % des visages[39]. C'est précisément pour contrer ce type d'utilisations que le corps médical doit s'approprier ces outils et s'impliquer dans leur développement.

Autre danger : le détournement de l'intelligence artificielle pour tromper volontairement les futurs systèmes médicaux qui l'utiliseront. À l'avenir, le *Deep Learning* sera massivement utilisé en radiologie ou en anatomopathologie. Ainsi, on peut imaginer qu'un intervenant tiers utilise des réseaux antagonistes génératifs (*Generative Adversarial Network*, voir annexes) afin de tromper automatiquement ces systèmes, sans que personne s'en aperçoive. En entraînant une IA dédiée, il est possible de générer des images médicales qui semblent tout à fait normales à l'œil humain, mais autorisent n'importe quel résultat lorsqu'elles sont interprétées automatiquement par *Deep Learning*[40]. Imaginez une radiographie de thorax de face tout à fait normale à l'œil nu, mais qui contiendrait les informations numériques nécessaires pour tromper un système de *Deep Learning* automatique afin de lui faire croire qu'il existe une anomalie sur la radio alors qu'il n'y en a pas ou inversement. Ces risques de manipulation et de fraude devront être contrés par des méthodes dédiées.

Enfin, il faut particulièrement veiller aux fausses promesses. Une équipe de Cambridge a récemment publié dans la revue *Nature Machine Intelligence* une analyse de la qualité scientifique des études publiées dans le domaine de l'IA appliquée au diagnostic et au pronostic du Covid-19 à partir de radiographies du thorax et de scanners[41]. L'équipe a ainsi recensé plus de deux mille études et sélectionné les soixante meilleures pour analyse. Il ressort de ce travail qu'aucune ne présentait les critères de qualité suffisants pour envisager d'utiliser un jour la méthode décrite en routine clinique : soit la méthode informatique n'était pas suffisamment documentée, soit elle ne suivait pas les recommandations de bonnes pratiques de développement, soit l'IA n'avait pas été testée sur une cohorte de validation externe. Il existe donc un hiatus énorme entre la masse des études publiées et le nombre d'études de qualité.

Une équipe du MIT a également évalué la qualité scientifique de 511 études publiées portant sur le *Machine Learning* appliqué à la médecine de façon plus générale[42]. Ils ont ainsi mis en évidence que les études publiées étaient de moins bonne qualité que dans d'autres domaines d'application de l'IA, notamment parce que les bases de données utilisées et le code informatique créé – afin de pouvoir les reproduire et vérifier leur authenticité – n'étaient pas accessibles. Il est donc crucial de créer et d'adopter des recommandations pour développer les IA de demain dans les meilleures conditions.

5

Secret médical et intelligence artificielle

La création d'algorithmes à partir de données de santé a déjà été à l'origine de différentes problématiques liées au respect du secret médical. Les exemples se sont récemment multipliés, parmi lesquels le projet *Nightingale* de Google[43]. Puisque le principe même de l'IA est d'utiliser des données afin d'entraîner un algorithme informatique, comment pouvons-nous assurer leur confidentialité et protéger le secret médical que le médecin s'est engagé à respecter lorsqu'il a prêté le serment d'Hippocrate ? Au-delà du cadre réglementaire, quelles techniques pourrions-nous utiliser afin de ne pas freiner le développement de l'IA, tout en garantissant le secret ?

La première technique est bien évidemment l'anonymisation complète des données médicales. Cette tâche qui peut sembler aller de soi est en réalité bien plus complexe qu'il n'y paraît. En effet, en anonymisant de façon exhaustive les données, on risque de perdre certaines informations qui pourraient être des facteurs pronostiques comme le sexe, le lieu de résidence, la date de naissance et donc l'âge du patient. II existe déjà des méthodologies permettant de générer des données de façon artificielle afin de remplacer la donnée réelle et donc d'assurer une bonne anonymisation. Leur utilisation pourrait cependant introduire des biais cachés liés aux algorithmes de synthèse eux-mêmes. D'un autre côté, si le nom du patient est simplement retiré du dossier médical, il restera toujours relativement aisé de l'identifier à partir de certaines particularités et de certains motifs de caractéristiques qui lui sont propres. Il existe donc un juste milieu entre ces deux approches très difficile à définir.

La seconde technique consisterait à entraîner un réseau antagoniste génératif à partir d'une cohorte réelle de patients pour créer une nouvelle cohorte totalement artificielle, et donc anonyme. Cette nouvelle cohorte synthétique contiendrait les mêmes spécificités et la même distribution statistique des caractéristiques que la cohorte réelle. Si cette solution apparaît séduisante, elle pose tout de même la question de l'acceptabilité d'une IA médicale entraînée et validée sur des données de synthèse et non réelles. Comment nous assurer des résultats d'une telle IA évaluée sur de fausses données ? Comment lui faire confiance ? Et au final, comment envisager son adoption par le personnel soignant ?

Une troisième solution serait d'utiliser l'apprentissage fédéré (*federated learning*). Cette méthode ne revient pas à rassembler les données d'un grand nombre de patients dans une structure unique pour entraîner un algorithme, mais au contraire à « exporter » l'entraînement de l'algorithme lui-même sur

chaque patient ou sur chaque cohorte afin qu'il apprenne de façon décentralisée. Une fois l'apprentissage réalisé, chacun des petits modèles est rapatrié pour recréer un modèle général optimal. Cette méthode permet de ne pas agréger les données de plusieurs centres dans un endroit unique et limite le risque de rupture du secret médical. La start-up française Owkin travaille particulièrement sur ces méthodologies innovantes. D'autres méthodes sont également en cours d'évaluation telles que l'*encrypted Machine Learning* où l'algorithme apprend sur des données encryptées, et donc confidentielles.

6

Comment gérer et utiliser nos données ?

On sait depuis maintenant plusieurs années que les GAFA (Google, Amazon, Facebook, Apple) s'intéressent de près à la santé. Aux États-Unis, la santé représente un marché porteur incontournable où les dépenses de soins correspondent à 20 % du PIB (produit intérieur brut), soit 3,5 trillions de dollars ! Dans ce contexte, Google avait débuté le projet *Nightingale*, en partenariat avec Ascension, le second plus grand groupe hospitalier des États-Unis (2 600 structures de soins). Les dossiers médicaux des patients traités dans les hôpitaux d'Ascension avait donc été transférés vers les infrastructures de Google. Mais les modalités de transfert et d'exploitation de ces données, apparemment pas suffisamment anonymisées, ont contribué à alimenter la polémique et les craintes sur leur utilisation ultérieure et les dérives éventuelles. Le fait que les patients concernés n'aient pas été informés de l'utilisation de leurs données a également été perçu comme problématique. Aux États-Unis, de tels accords sont en réalité assez fréquemment passés dans le cadre de BAA (*Business Associate Agreement*), si le partenariat est établi pour le bénéfice du patient et en conformité avec la loi HIPAA (*Health Insurance Portability and Accountability Act*). Mais qu'en est-il en France ? Un tel partenariat serait-il possible ?

Notre pays dispose d'une quantité de données de santé centralisées parmi les plus importantes au monde, notamment en raison de la structure de son système de remboursement des soins. Il existe donc d'une part une responsabilité des organismes pour protéger ces données, mais probablement aussi une responsabilité pour les exploiter et faire ainsi progresser la recherche médicale, pour le bénéfice direct des patients. Ne pas exploiter ces données serait clairement dommageable à l'heure du *big data* et de l'intelligence artificielle. Un équilibre fin doit être trouvé entre restriction de l'accès aux données de soins, et absence totale de contrôle.

Depuis le 25 mai 2018, le *Règlement général sur la protection des données* (RGPD) est entré en vigueur dans l'ensemble de l'Union européenne. Dans ce cadre, il appartient aux organismes (hôpitaux, cliniques, assurance maladie, mutuelles) de cartographier leur traitement des données de santé, de gérer les risques, d'organiser les processus internes de gestion et de documenter leur conformité à la loi. Concrètement, le RGPD doit simplifier les démarches et responsabiliser les acteurs. Plusieurs exemples emblématiques de structuration de l'accès aux données de santé se mettent en place en France. Parmi elles, le *Health Data Hub* est déployé par le ministère

des Solidarités et de la Santé : il a pour objectif de favoriser l'utilisation des données de santé et devra permettre à terme le développement de nouvelles techniques, en articulation avec les instituts interdisciplinaires d'intelligence artificielle (3IA). Plusieurs projets pilotes ont déjà été sélectionnés (cancer, cardiologie, neurologie) et d'autres projets sont à venir.

À l'échelle régionale, l'Assistance publique – Hôpitaux de Paris a également structuré les informations de ses trente-neuf hôpitaux dans un gigantesque entrepôt de données de santé regroupant environ dix millions de patients et vingt-cinq millions d'examens d'imagerie. L'accès et l'exploitation de ces données sont aussi strictement encadrés, notamment par l'évaluation préalable de chaque projet par un conseil scientifique et éthique. On peut donc se réjouir des avancées concrètes qui ont eu lieu en France car elles vont permettre de faire avancer la médecine dans de nombreux domaines, au sein d'un cadre réglementaire protecteur.

La réflexion sur les données de santé doit cependant se poursuivre : il est par exemple nécessaire de déterminer si elles appartiennent au patient lui-même ou à l'entité qui les a générées. Si ces informations appartiennent avant tout au patient, ce que semblent suggérer les évolutions récentes du droit et de la jurisprudence, alors un nouveau paradigme de gestion et d'utilisation de ces données devra être défini. Les patients pourront-ils disposer de leurs données et les vendre pour être rémunérés ? Ou au contraire, la vente de ces éléments sera-t-elle interdite ? Pourront-ils en faire don à la recherche ? On le voit, l'analogie avec le statut légal du corps humain est tentante. Mais est-elle pertinente ? Dans le droit français, le corps humain est inviolable et indisponible. Il est ainsi interdit de porter atteinte à l'intégrité corporelle d'une personne sans son consentement et toute convention à titre onéreux portant sur le corps et ses produits (organes, gamètes) est considérée comme nulle. Les données doivent-elles être considérées comme un produit du corps humain ? Sont-elles inviolables et indisponibles ? Cela limiterait sévèrement leur utilisation et donc le développement de l'IA médicale en France.

7

Faut-il réguler l'intelligence artificielle ?

La régulation de l'IA doit répondre à deux contraintes *a priori* antagonistes : protéger les patients d'IA inefficaces et ne pas freiner l'innovation. Pour ce faire, le concept de « garantie humaine » a été développé par l'initiative Ethik-IA, repris dans le rapport Villani sur l'intelligence artificielle sorti en 2018 et enfin dans l'avis 129 émis par le Comité consultatif national d'éthique, dans le cadre de la révision de la loi de bioéthique. Le principe global est de mettre en place une supervision humaine de l'IA à travers le déploiement de collèges de garantie associant médecins, paramédicaux et représentants des usagers. Ce collège effectuerait une révision *a posteriori* des dossiers médicaux, afin de vérifier l'efficacité de l'IA.

Nous l'avons vu, l'IA excelle dans les tâches de perception : elle est en train de changer la radiologie diagnostique et l'anatomopathologie, entre autres. Dans ces domaines, l'utilisation de l'IA sera vérifiable et ses résultats pourront être évalués par un humain. Le concept de « garantie humaine » pour ce type de tâches est donc tout à fait pertinent... jusqu'à un certain point.

La garantie humaine semble en effet présupposer que l'humain sera toujours meilleur que l'IA et qu'il sera donc toujours à même de déterminer si l'IA a été efficace. En tenant compte des évolutions techniques dans le domaine et des résultats de certaines études comparant les performances diagnostiques d'une IA à celles d'un médecin, il semble peu probable que les humains restent encore longtemps à même de juger une IA. Que se passera-t-il si un ou des humains « mauvais » censurent une IA qui en réalité a raison ? Comment savoir si le collège humain a raison ou tort ? Le patient doit-il accepter de ne pas bénéficier d'une IA superperformante sur le principe de la garantie humaine ?

Dans d'autres domaines, les résultats fournis par une IA ne seront tout simplement pas vérifiables : prédiction de l'apparition d'une maladie cinq ans à l'avance, prédiction du risque de décès, analyse génomique. Le médecin et donc le patient n'auront d'autre choix que de « faire confiance à l'IA ». Cela pose un véritable problème, car on ne sait pas encore comment valider ce type d'IA. Dans ce contexte, faut-il réguler l'utilisation de l'IA en médecine, comme on contrôle la mise sur le marché des médicaments et des dispositifs médicaux ? Si oui, quelles études cliniques faire, quelles méthodes utiliser ? Faut-il imiter le protocole de mise sur le marché des médicaments pour vérifier leur innocuité et leur efficacité, en traitant un groupe de patients selon les recommandations d'une IA et en comparant les résultats à ceux d'un autre

groupe traité par un médecin ? Ces questions complexes restent ouvertes.

Comment évaluerons-nous une intelligence artificielle utilisée pour prédire le devenir d'une maladie, afin de personnaliser les soins ? Par exemple pour prédire la guérison ou la récidive d'un patient atteint d'un cancer. Si l'IA prédit malheureusement une récidive, et que nous intensifions les traitements, comment savoir si la guérison est due aux traitements ou à une erreur de prédiction ?

Où en sommes-nous véritablement ? Aux États-Unis, en un an et demi, la FDA (autorité qui régule la mise sur le marché américain des médicaments) a autorisé la commercialisation de treize solutions médicales s'appuyant partiellement ou totalement sur l'IA dans le diagnostic d'hémorragie cérébrale, d'accident vasculaire cérébral et la détection du cancer du poumon ou du foie. Ces systèmes sont actuellement vendus et peuvent être utilisés en routine pour diagnostiquer et donc traiter des patients. Malheureusement, un seul de ces dispositifs a fait l'objet d'une étude clinique rigoureuse et les douze autres sont vendus sans aucune validation préalable. Les diagnostics et recommandations faites par ces systèmes d'IA pourraient se révéler faux, impossibles à vérifier et donc délétères pour les patients. En Europe et en France, l'AME (Agence européenne des médicaments) et l'ANSM (Agence nationale de sécurité du médicament et des produits de santé) utilisent des critères d'homologation plus stricts, ce qui tend à mieux protéger les patients.

Il est donc important de prendre du recul sur l'utilisation de l'IA en médecine et de ne pas brûler les étapes afin de ne pas exposer la population à d'éventuelles erreurs, impossibles à détecter. Dans le cas d'une erreur liée à l'utilisation d'une IA, la question de la responsabilité finale se pose. Est-ce le médecin qui a pris la décision ? La faute revient-elle à l'entreprise qui commercialise l'IA, ou encore à l'hôpital qui a déployé le produit ? Une étude récente, publiée dans le *Journal of Informatics in Health and Biomedicine* (JAMIA), a évalué le point de vue des médecins et du public[44]. Il apparaît que le public est plus à même de considérer que la faute, et donc la responsabilité, incombe au médecin. En effet, 66 % des personnes interrogées considèrent que les médecins sont responsables en cas d'erreur de l'IA, et 57 % des médecins interrogés partagent la même opinion. Ceux-ci jugent également que les entreprises commercialisant l'IA (44 %) et les hôpitaux (29 %) doivent être tenus pour responsables.

En parallèle, la compétition internationale, États-Unis et Chine en tête, pousse à aller toujours plus vite. Les autorités sanitaires et le public sont donc devant un choix cornélien : réguler fortement l'utilisation de l'IA en santé et prendre un retard considérable, ou ouvrir grand les portes du marché du soin aux solutions d'IA, au risque de favoriser l'émergence de systèmes défectueux, voire dangereux. La garantie humaine est une première réponse, au moins pour les IA de perception et les quelques années à venir, mais après,

comment faire ?

8

Comment seront créées les intelligences artificielles de demain ?

Étant donné le rôle important qu'est appelée à jouer l'intelligence artificielle en médecine, et considérant les difficultés de validation d'une intelligence artificielle pour la prise en charge des patients, des recommandations précises sont nécessaires afin de s'assurer de la bonne qualité des outils que nous utiliserons à l'avenir.

La recherche médicale dans les années 1970 et 1980 a rencontré de nombreux problèmes liés à la qualité médiocre des méthodologies utilisées pour valider de nouveaux traitements. En réponse à cette problématique, la communauté des chercheurs a développé des recommandations en fonction du type d'étude et de traitement à évaluer, afin d'assurer un niveau de qualité acceptable. En 1996, le groupe CONSORT a ainsi proposé une standardisation de la méthodologie d'un essai clinique.

Le développement de ces recommandations est complexe et fait souvent appel à des équipes d'experts pluridisciplinaires venus d'horizons divers : universitaires, chercheurs, praticiens, responsables politiques ou encore groupes de patients. Les recommandations produites correspondent en général à des check-lists de quinze à trente items qui doivent figurer dans la rédaction de l'étude. La particularité du développement d'une intelligence artificielle vient des différentes étapes à passer avant de pouvoir être utilisée.

La première de ces étapes est le développement et la validation *in silico*, c'est-à-dire uniquement dans l'ordinateur, de façon virtuelle, d'une IA. Plusieurs recommandations à ce stade, globalement redondantes : *PRIME*, *TRIPOD-ML*, *MINIMAR*, *CLAIM*, *MI-CLAIM* et *STARD-AI*. Il s'agit des étapes nécessaires exposées dans l'annexe « Comment créer une intelligence artificielle ? »[45].

La seconde étape est l'évaluation clinique intermédiaire. Si le développement *in silico* d'une IA correspond à une preuve de concept, il faut pouvoir vérifier que cette preuve est valable avant de lancer des essais cliniques à large échelle. Cette étape intermédiaire sert donc à évaluer, par de petites expériences cliniques, l'ergonomie, la justesse de l'IA et l'influence des facteurs humains sur son utilisation. Le retour d'expérience permettra de prototyper de façon agile et rapide, dans un environnement clinique simulé, l'application potentielle de la nouvelle IA. Concrètement, cette étape correspond aux phases 1 et 2 nécessaires au développement d'un nouveau

médicament, durant lesquelles on évalue ses effets secondaires et sa possible efficacité. La recommandation *DECIDE-AI* a récemment été publiée pour cette étape.

La troisième étape, sans doute la plus compliquée à mettre en place, est la réalisation d'essais prospectifs multicentriques. Les essais cliniques randomisés sont considérés comme le standard obligatoire avant l'utilisation et la commercialisation de tout nouveau médicament ou de tout nouveau dispositif médical. Mettre en place ce type d'essais pour évaluer une IA est particulièrement complexe. La communauté des chercheurs a peu d'expérience dans ce nouveau domaine et ne peut reprendre telles quelles les méthodes applicables aux médicaments. Deux recommandations sont actuellement développées : *SPIRIT-AI* et *CONSORT-AI*. Elles tiennent compte des problématiques spécifiques à l'IA, comme les critères d'inclusion et d'exclusion des patients qui doivent également correspondre à ceux utilisés dans la base de données de développement initial de l'IA. Autre problème : alors qu'on évalue souvent un nouveau médicament par rapport à un placebo ou par rapport à un ancien médicament, comment évaluer une IA qui n'a pas d'équivalent ou qui exécute une tâche impossible pour un humain ? De même, comment évaluer à un instant t un modèle d'intelligence artificielle qui se mettrait continuellement à jour avec de nouvelles données ?

D'autres recommandations permettent de guider les hôpitaux ou les cliniques dans l'achat de nouvelles solutions. Ces recommandations, appelées ECLAIR, évaluent la pertinence d'un système d'IA par rapport à la structure qui souhaiterait éventuellement en faire l'acquisition.

9

Qui sera le leader en intelligence artificielle ?

Selon le rapport *AI Index* du Stanford Institute for Human-Centered Artificial Intelligence, les investissements privés en IA se sont élevés à plus de quatre-vingt-treize milliards de dollars dans le monde en 2021[46]. Les plus gros investissements ont été réalisés aux États-Unis, en Chine et en Europe. Les start-up récoltent la majeure partie de ces investissements, avec plus de trente-sept milliards de dollars, en très nette augmentation par rapport aux 1,3 milliard levés en 2010. Tous domaines confondus, le marché de l'IA représentait 156 milliards de dollars en 2020, avec une augmentation de plus de 13 % par rapport à 2019.

Cette inflation se retrouve aussi dans la recherche en IA, puisque le nombre d'articles publiés est passé de 52 000 en 2000 à environ 400 000 par an aujourd'hui. La Chine publie davantage que les États-Unis, leader incontesté du domaine jusqu'en 2016. Cependant, les articles publiés par la Chine sont globalement bien moins cités que ceux des Américains. Cela reflète-t-il la qualité et la fiabilité des études ou la façon dont elles sont perçues par la communauté scientifique ? Dans ce classement, la France est sixième, après les États-Unis, la Chine, le Royaume-Uni, l'Inde et l'Allemagne. Les différentes estimations placent encore les États-Unis en première position pour de nombreuses années à venir. Après la compétition spatiale entre États-Unis et Russie, c'est la course à l'hégémonie en IA entre les Américains et la Chine.

Aux États-Unis, la National Science Foundation a annoncé en 2021 la création de cinq instituts dédiés à l'IA, chacun géré par une université dotée de vingt millions de dollars sur cinq ans. Avec d'autres agences américaines, notamment dans le domaine de la défense et de la santé, le gouvernement américain a investi cinq milliards de dollars dans la recherche sur l'IA en 2020.

La Commission européenne estime que les investissements publics et privés en IA à travers l'Europe avaient atteint vingt milliards d'euros en 2020. Dans la réglementation de l'IA proposée par l'Union européenne, un texte de quatre-vingt-une pages, le mot santé (*healthcare*) figure à plusieurs reprises, mais aucune section spécifique n'y est dédiée. Le terme médecine (*medicine*) n'est pas mentionné. Faut-il se réjouir du fait que l'UE ne cherche pas à réguler fortement l'IA en médecine ou doit-on plutôt y voir un oubli fâcheux ? On ne peut croire que l'UE n'ait pas pleinement pris conscience de l'importance de cette thématique. Quoi qu'il en soit, ce nouveau cadre légal

devra créer un environnement pour à la fois protéger les citoyens européens et ne pas freiner le développement de l'IA.

En France, les ministères de l'Économie et des Armées ont lancé en 2019 le plan *Deep Tech* au budget de 2,3 milliards d'euros. Dans le cadre de la stratégie nationale pour l'intelligence artificielle annoncée par le président de la République en mars 2018, quatre instituts interdisciplinaires d'intelligence artificielle (3IA) ont été créés à Grenoble, Nice, Paris et Toulouse. Le budget de ces instituts s'élève à 225 millions d'euros pour quatre ans. On le voit, les gouvernements des principales puissances du monde ont parfaitement compris les enjeux stratégiques et économiques de l'IA. Au-delà des financements, ce qui conditionnera notre capacité à devenir leader dans le domaine est d'une part la facilité d'accès aux données et d'autre part notre capacité à garder et à faire venir les meilleurs cerveaux.

Une enquête réalisée en 2021 auprès de 524 chercheurs en IA du monde a montré quels pays les attiraient le plus. Sans surprise, les États-Unis arrivent en tête (59 %), suivis du Royaume-Uni (35 %), du Canada (28 %), de la Suisse (25 %) et de la France (16 %). Les raisons invoquées incluent les opportunités professionnelles, la qualité de vie, la culture, le climat politique et les relations personnelles{47}.

10

2031 : une semaine, du dépistage au traitement d'un cancer du poumon

Reprenons l'exemple de M^me^ Emma Langlois, 53 ans, évoqué au début du livre. En 2031, Apple a lancé depuis trois ans son programme de suivi de la santé en temps réel, *Apple Life*. Les enfants de M^me^ Langlois lui ont offert à Noël dernier une montre connectée et un abonnement d'un an au programme.

En 2027, Apple a publié dans le *New England Journal of Medicine* une étude démontrant l'intérêt du *quantified-self*, soit le fait de mesurer un grand nombre de paramètres de santé à l'échelle individuelle tels que le poids, l'activité physique, la fréquence cardiaque, la saturation en oxygène du sang, la glycémie et l'état mental. Le système comprend un coach artificiel très réaliste, entièrement adapté à la personnalité de l'utilisateur, qui passe des appels audio et vidéo et envoie des messages pour le motiver à améliorer ses comportements, son alimentation... et réduire le temps qu'il passe devant un écran. Les résultats sont impressionnants : grâce à cette approche, on est capable d'améliorer l'état de santé des participants et de détecter très rapidement une anomalie nécessitant une consultation. En comparant de façon prospective la santé et le bien-être de deux groupes de 35 000 personnes, l'un avec un suivi « classique » et l'autre avec un suivi « quantifié » par *Apple Life*, les chercheurs ont montré une diminution du surpoids et de l'obésité, une augmentation de l'activité physique et une diminution du diabète. La méthode permet aussi de détecter très précocement l'apparition de symptômes liés à certaines maladies, dont le cancer, permettant des traitements plus rapides, plus efficaces et avec moins de séquelles en raison des stades moins avancés ainsi diagnostiqués. Les économies de santé sont également majeures. Devant de tels résultats, la FDA approuve le programme *Apple Life* trois mois après la publication, l'European Medicines Agency lui emboîte le pas quinze jours plus tard. En France, une procédure d'autorisation de mise sur le marché accélérée est réalisée et le dispositif devient remboursé par la Caisse primaire d'assurance maladie. Ne pas utiliser cette méthode est rapidement considéré comme une véritable perte de chance pour les patients.

En parallèle, des voix s'élèvent pour dénoncer l'irruption d'un acteur du numérique dans la santé et la monétisation du dispositif. D'autres relèvent que les données collectées par Apple dans le monde entier risquent d'être détournées de leur bon usage : alors qu'elles sont censées permettre d'améliorer sans cesse l'algorithme de détection et les recommandations de santé que le programme préconise à chaque participant, certains craignent

qu'Apple ne s'en serve ou ne les revende pour cibler des publicités. D'autres enfin s'inquiètent de ce nouveau dispositif de surveillance massive et de ses conséquences sur les libertés individuelles.

Pour Apple, qui n'a plus sorti de produit révolutionnaire depuis l'iPhone en 2007, l'enjeu commercial est de taille, surtout après l'échec relatif de son casque de réalité virtuelle en 2023. La campagne marketing est savamment orchestrée : « *Live Different* » reprend l'histoire de son fondateur Steve Jobs, emporté à 56 ans par un cancer du pancréas diagnostiqué tardivement. Le produit est un succès historique. La première année, les ventes et les abonnements d'*Apple Life* représentent plus de 10 % des revenus de la firme.

Il est 7 h 40 ce lundi matin de juin 2031 lorsque M^me^ Langlois reçoit une notification sur sa montre : « Votre activité physique a été réduite de 10 % sur le dernier mois. Votre fréquence cardiaque a augmenté de 7 % et la saturation en oxygène de votre sang a diminué de 3 %. Veuillez consulter votre médecin. » Elle reçoit en même temps, ainsi que son médecin, un rapport qui détaille les anomalies constatées : « Ces anomalies sont corrélées à un risque accru de cancer du poumon. *Apple Life* vous recommande de consulter votre médecin traitant au plus vite et de passer un scanner thoracique. » Emma est assez étonnée par la froideur statistique du message et sent déjà l'angoisse monter en elle. Elle prend donc rendez-vous chez son médecin l'après-midi. Les délais sont habituellement beaucoup plus longs, mais le praticien a des créneaux réservés pour les patients *Apple Life*. Ce principe est d'ailleurs très critiqué car beaucoup considèrent qu'il participe à l'instauration d'une médecine à deux vitesses.

Le D^r^ Michel reçoit M^me^ Langlois avec la demi-heure de retard habituelle. Même en 2031, l'IA ne peut pas tout régler ! Le médecin parcourt le rapport et l'explique à Emma en détail. Il est d'accord avec les conclusions et la conduite à tenir : il faut faire un scanner thoracique sans tarder.

Depuis l'avènement en France de la lecture automatisée des scanners par IA, les délais de rendez-vous ont été divisés par 10. De trois semaines minimum auparavant, nous sommes passés à trois jours. Emma passe donc son examen le jeudi. Cinq cents coupes sont réalisées en quelques minutes et interprétées instantanément : l'IA a détecté un nodule de 15 millimètres dans le lobe supérieur du poumon droit. Le radiologue, le D^r^ Remy, interprète à son tour le scanner et confirme les résultats de l'IA. Il voit Emma en consultation pour lui expliquer directement le résultat. Le métier de radiologue, bien loin d'avoir disparu avec l'IA, a profondément changé : il fait maintenant des consultations, voit ses patients, les suit. Son activité s'est recentrée sur l'humain. Le D^r^ Rémy explique à M^me^ Langlois que le bilan doit être complété par une biopsie transthoracique du nodule, une fibroscopie pulmonaire, des explorations fonctionnelles respiratoires (EFR), la recherche d'ADN tumoral circulant, une IRM cérébrale et un scanner du corps entier ou un TEP-scan. Les deux derniers examens peuvent être réalisés dans la foulée. Il lui propose un rendez-vous le lendemain, vendredi, pour effectuer la biopsie

transthoracique sous scanner avec une anesthésie locale. Le reste des examens devra être fait rapidement à l'hôpital.

Le système national de gestion des maladies graves, initié en France en 2029, a déjà attribué une priorité haute à M^me^ Langlois. Un rendez-vous le lundi suivant avec un oncologue spécialisé dans le cancer du poumon est déjà programmé. Ce système, qui repose sur une IA de gestion de rendez-vous et de flux, permet d'optimiser les soins pour limiter les délais de prise en charge et les déplacements des patients. Il prend en compte le diagnostic, surveille et coordonne la disponibilité des rendez-vous sur la région dont dépend le patient et tient compte de ses préférences. Depuis sa création par une équipe de l'INSERM, ce système a permis de réduire de 60 % les délais de diagnostic et de traitement, de 50 % les plages de rendez-vous inutilisées et de 20 % ceux non honorés par les patients. Ce système optimise le parcours de soins du patient mais tient aussi compte de son agenda. M^me^ Langlois passera donc ses derniers examens et aura rendez-vous avec son oncologue le même jour, dans le même hôpital. En tout et pour tout, de l'alerte *Apple Life* à la réalisation de l'ensemble du bilan, cinq jours se sont écoulés.

Entre la biopsie vendredi et le rendez-vous avec l'oncologue le lundi suivant, le prélèvement a été préparé en laboratoire par des robots puis analysé par une IA pour déterminer la nature de la tumeur et ses caractéristiques. Ce qui prenait auparavant une dizaine de jours est maintenant accompli en deux jours, validation par un anatomopathologiste comprise.

Le lundi, M^me^ Langlois se rend à ses rendez-vous. À peine a-t-elle franchi la porte de l'hôpital que sa présence est détectée par HIA, l'intelligence artificielle qui gère l'ensemble de l'hôpital. HIA dispose de caméras et de capteurs dans l'établissement, reconnaît les patients, les guide jusqu'à leur rendez-vous et les prévient en cas de retard. En parallèle, elle avertit le médecin de l'arrivée du patient, prépare le dossier médical adéquat et recommande la meilleure attitude à adopter en fonction de l'état émotionnel du patient.

Durant la consultation, la P^r^ Marte lui annonce progressivement la mauvaise nouvelle avec tact et mesure : elle a bien un cancer du poumon, un adénocarcinome, mais heureusement confiné au poumon, sans métastase car il a été détecté très tôt. Plusieurs traitements curatifs sont possibles : la chirurgie ou la radiothérapie stéréotaxique. M^me^ Langlois choisit la radiothérapie pour continuer d'exercer son activité professionnelle qu'elle ne peut et ne veut pas arrêter.

Le mercredi, M^me^ Langlois se rend dans le service de radiothérapie. Elle rencontre son oncologue radiothérapeute qui lui explique comment se déroulera le traitement, puis passe un scanner dosimétrique. Cet examen sert à calculer une dosimétrie et programmer la machine qui délivrera le traitement. En trente minutes, l'examen est réalisé, la tumeur contourée et la dosimétrie terminée. Ce qui pouvait prendre trois semaines auparavant dure à peine une

demi-heure. La séance de traitement a lieu l'après-midi même et M^me^ Langlois peut rentrer chez elle avant 18 heures.

Un suivi très rapproché est assuré par le *quantified-self*. Un algorithme surveillera dans le temps différents paramètres comme le poids, ou l'apparition d'une fatigue ou de symptômes : au moindre élément suspect, M^me^ Langlois reverra son médecin et passera un scanner à la recherche d'une récidive. Là encore, l'IA permettra d'optimiser son suivi et de diagnostiquer très précocement toute récidive pour pouvoir la traiter le plus tôt possible et augmenter ses chances de guérison.

11

2041 : une IA remporte trois prix Nobel

En 1903, Marie Curie fut la première femme à recevoir le prix Nobel de physique avec son mari Pierre Curie pour ses travaux sur les radiations. Huit ans plus tard, en 1911, elle devint aussi la seule femme à avoir obtenu deux prix Nobel lorsqu'elle fut récompensée d'un second prix, cette fois en chimie. Cent trente ans plus tard, GAIA (pour *General Artificial Intelligence for All*) devient la première entité non humaine à obtenir successivement le prix Nobel de physique, de chimie, puis de médecine pour ses travaux en physique quantique, en pharmaceutique et en oncogenèse.

GAIA est un système global, une intelligence artificielle générale, capable de traiter un nombre infini de tâches simultanément, sans limite dans ses domaines d'applications. Elle a été conçue par une équipe de scientifiques financés par un consortium international rassemblant des universitaires et des chercheurs du domaine privé. Ce projet pharaonique, doté de plusieurs dizaines de milliards de dollars, a longtemps été comparé au projet *Manhattan*, qui produisit la première bombe nucléaire, en raison de son ampleur, mais aussi des dangers potentiels qu'il représente. Ses détracteurs craignaient que l'IA échappe au contrôle humain et cherche à exterminer l'humanité.

Le 16 juillet 2038, dix ans après son lancement, les premiers modules de GAIA sont activés et mis en ligne. Le système apprend de façon exponentielle et apporte déjà des réponses là où les plus grandes équipes scientifiques butent. Dans le domaine médical, GAIA est capable de comprendre la physiopathologie d'un nombre croissant de maladies et ainsi de créer des traitements personnalisés. Elle peut déterminer les anomalies génétiques à l'origine du développement d'un cancer et synthétiser la molécule qui ciblera ces anomalies pour empêcher l'apparition d'une tumeur ou la guérir.

Elle développe des modules-filles qu'elle déploie dans toutes les spécialités et qui sont reliés les uns aux autres. GAIA peut aussi sélectionner les meilleurs embryons sur des critères morphologiques et génétiques. L'eugénisme devient une réalité et il est inenvisageable de ne pas concevoir avec l'aide de GAIA si l'on souhaite que son enfant ne soit pas discriminé pour le reste de sa vie. Dans l'Union européenne, le Parlement tente de légiférer et de limiter les pouvoirs de GAIA, en vain. Le système permet de dépister avec une grande précision l'ensemble des maladies et de les prévenir. Elle est capable à partir d'une prise de sang, de sueur ou d'une larme de prédire l'apparition d'une maladie jusqu'à cinq ans à l'avance. Si ce risque

dépasse 5 %, elle propose un médicament préventif et un suivi dans le réseau de cliniques *Amazon Healthcare*, réseau devenu monopolistique dans le monde entier et avec lequel GAIA a développé un partenariat exclusif qui assure son fonctionnement, sa croissance et sa gratuité.

Dans les cliniques Amazon, un médecin par spécialité proposée est toujours présent. Il surveille les performances de GAIA, quand bien même le système a dépassé les performances humaines depuis plusieurs mois, comme elle l'a elle-même démontré dans un ensemble de plusieurs centaines d'études prospectives randomisées menées dans le réseau de cliniques où elle s'est déployée. Meilleur diagnostic, meilleur traitement, meilleure chirurgie robotique, le système ne connaît aucune limite. Les taux de guérison s'envolent, les séquelles des traitements sont de plus en plus rares. Les erreurs médicales et les aléas thérapeutiques deviennent quasi inexistants.

Dès les premiers mois de GAIA, de nombreuses voix s'étaient élevées contre l'utilisation du système dans le domaine médical : les médecins, tout d'abord, craignaient une déshumanisation de la médecine et une dégradation des soins. Les patients eux-mêmes appréhendaient de ne plus avoir d'interlocuteurs humains ni ce lien si particulier à leurs soignants. Mais rapidement, devant les performances exceptionnelles de GAIA, son empathie parfois beaucoup plus grande que celle d'un praticien, toutes ces craintes ont été balayées et, petit à petit, chacun a fait le choix de désigner GAIA comme son médecin traitant, y compris les médecins, qui auparavant se suivaient et se traitaient eux-mêmes. Plus de pénurie de médecins, plus de déserts médicaux, GAIA est partout.

En 2040, GAIA ouvre sa première faculté de médecine pour former les médecins humains qui devront la surveiller. Le cycle est désormais complet : alors que les humains ont cherché à dresser des machines contre la maladie pendant des siècles, c'est maintenant la machine qui apprend aux humains comment ils devront se soigner.

12

Quel avenir pour la médecine ?

En 2022 l'espérance de vie des hommes n'a jamais été aussi longue, grâce aux progrès de l'hygiène et de la médecine. Le dépistage, le diagnostic et les traitements des maladies permettent de guérir des pathologies qui étaient encore mortelles il y a quelques années et de rendre chroniques celles que l'on ne parvient pas à guérir. Pour autant, la lourdeur du parcours de soins et du système de santé, que l'on se trouve en France, quelque part en Europe, aux États-Unis ou en Chine, reste majeure. Les procédures de diagnostic ne sont pas infaillibles et les traitements, pas toujours taillés sur mesure pour chaque patient, entraînent parfois des effets secondaires et des séquelles importantes.

En parallèle, une nouvelle technologie, reposant sur l'analyse de données massives par des ordinateurs toujours plus puissants, permet d'obtenir des systèmes capables de réaliser des diagnostics, des prédictions, des traitements et bientôt des interventions chirurgicales de façon au moins aussi bonne que les humains dans certains domaines et très probablement bientôt meilleure. La majeure partie de ces systèmes n'est pas encore validée pour une utilisation clinique en routine. En l'état actuel des choses, il est tout à fait possible que dans cinq à dix ans nous en soyons toujours à évaluer ces systèmes de façon rétrospective, donc insuffisante, si la recherche médicale ne passe pas à l'étape suivante, c'est-à-dire à l'évaluation prospective du bénéfice clinique réel de ce type de systèmes. Nous pourrions tout à fait nous retrouver dans un nouvel hiver de l'intelligence artificielle, comme cette discipline en a déjà connu par le passé. Les chercheurs, les soignants et les patients, déçus par les promesses non tenues, pourraient non seulement se détourner de ce type de méthodes, mais aussi les rejeter en bloc.

Ma conviction est qu'en l'absence d'une récession économique majeure et mondiale la compétition scientifique et industrielle internationale favorisera l'émergence de systèmes d'IA parfaitement validés cliniquement et bénéfiques pour le patient. Une partie significative des procédures diagnostiques par imagerie et par anatomopathologie sera réalisée automatiquement à très haut débit par des machines. Ces procédures resteront, du moins au début, validées manuellement par un humain. À terme cependant, et plus particulièrement lorsque les performances de l'intelligence artificielle auront très largement dépassé celles de l'humain, le concept même d'une validation ou d'une vérification par un médecin n'aura plus aucun sens. Nous aurons atteint un point de non-retour. Quel rôle joueront les médecins chargés jusqu'alors de ce type de procédures ? Comment assurerons-nous le contrôle

qualité de ces systèmes si nous ne sommes pas capables de comprendre les raisons pour lesquelles ils fonctionnent bien et si leurs performances dépassent les nôtres ? C'est à cela notamment que doivent être consacrés les efforts de recherche.

D'ici quinze à vingt ans, c'est le domaine du thérapeutique qui sera révolutionné par l'IA. Elle commencera par recommander les dispositifs et médicaments les plus adaptés à chaque maladie, mais aussi à chaque patient, de façon bien plus précise qu'un cerveau humain peut le concevoir. Conjuguée à la robotique, elle a effectué des interventions chirurgicales irréprochables sur le plan technique. Dans le même temps, ou quelque temps après, elle sera capable de mener une psychothérapie en trouvant les meilleurs traitements pour soigner les troubles mentaux.

Un peu plus tard, peut-être dans trente à quarante ans, l'IA sera à même de donner une consultation médicale parfaite du point de vue technique, mais aussi et surtout du point de vue « humain ». Avant même que le patient n'entre dans le box de consultation, elle connaîtra sa personnalité, elle saura trouver les mots justes pour lui parler, ce qui est sans doute le plus difficile. Elle sera alors sans doute bien plus empathique que la majorité des médecins humains et évidemment bien plus efficace.

Il semble évident qu'au début une majorité de patients refusera d'être prise en charge selon ces modalités et préférera toujours le contact avec une personne. Il est vraisemblable qu'avec les années et l'acceptabilité les patients souhaiteront être pris en charge par une IA, car cela sera à la fois plus agréable et plus efficace qu'une prise en charge par un médecin humain. *Plus humain qu'humain*, c'est le slogan de la Tyrell Corporation, l'entreprise qui commercialise les *replicants* dans *Blade Runner*, film sorti en 1982. Les androïdes du film sont d'abord perçus par les spectateurs comme froids et inhumains. Mais c'est finalement l'un d'entre eux qui sauvera la vie de celui qui cherche à les éliminer. Dans ce cas, qui est le plus humain ? Il est tout à fait possible que l'IA appliquée à la médecine atteigne un jour ce degré d'« humanité » nécessaire à la relation médecin-patient.

Admettons un instant que ce scénario digne des meilleurs (ou des pires) films de science-fiction se réalise. Nous sommes alors en 2041 et tous les patients dans les pays développés sont traités par une IA sous une forme ou une autre. Aurons-nous encore besoin de former des médecins ? Et si oui, comment pourrons-nous les former si l'IA réalise elle-même la majorité des tâches médicales ? Allons encore plus loin : puisqu'une IA est toujours créée à partir de données du passé, comment ferons-nous évoluer le savoir médical dans un contexte où toutes les pratiques, aussi automatiques et parfaites soient-elles, sont fondées sur le passé ? Même en implémentant des systèmes dits d'apprentissage continu où l'IA poursuit son apprentissage sur chaque nouveau cas, la question reste pertinente. Si l'IA continue d'apprendre de façon continue à partir de chaque patient, comment pouvons-nous nous

assurer que son efficacité reste correcte ? L'IA et la robotique en viendront-elles à former elles-mêmes les générations suivantes de robots et d'IA, hors de toute intervention, tout contrôle humain ?

À la lecture de ces lignes, on peut penser que ce type de scénario ne se réalisera jamais. Mais revenons alors au milieu des années 1990, lorsque Internet n'existait pas. Si vous faites partie d'une génération qui a connu un monde sans Internet, vous savez qu'en vingt-cinq ans ce réseau informatique a totalement révolutionné le monde. Comment pouvions-nous deviner en 1995 les changements radicaux qu'Internet occasionnerait ? Deuxième exemple : alors qu'aujourd'hui nous ne pouvons imaginer une journée sans notre smartphone, rappelez-vous à quoi ressemblait votre journée avant le 9 janvier 2007, jour de la présentation du premier iPhone par Steve Jobs.

Il y a de nombreuses raisons d'espérer en l'évolution de la médecine à l'âge de l'IA, mais tout autant de la craindre. De nombreuses questions restent sans réponse. Il nous appartiendra donc de rester optimistes et de ne pas freiner le développement de l'IA, mais aussi de nous montrer prudents dans les processus que nous utiliserons pour l'évaluer et l'utiliser. Au-delà de ces thématiques, nous devons également veiller à ce que le développement de l'IA ne favorise pas une médecine à deux vitesses où seuls les patients les plus aisés pourront bénéficier des toutes dernières technologies.

Annexes

Comment créer une intelligence artificielle ?

Afin de démystifier ce qu'est actuellement une *intelligence artificielle*, je vous propose de vous expliquer les principales étapes qui permettent d'en créer une. La création de modèle d'IA se déroule généralement en deux grandes étapes : le développement et la validation. Un premier ensemble de données est utilisé pour déterminer un ou plusieurs facteurs prédictifs. Lorsque ces facteurs ont été définis, ils sont évalués et validés sur un second ensemble de données différent du premier. De façon un peu plus détaillée, on suit les étapes suivantes :

1. Sélectionner et nettoyer les données.
2. Sélectionner la meilleure méthode d'analyse, adaptée à l'objectif du modèle et aux données de départ.
3. Valider la généralisation du modèle avec un ensemble de données externes.
4. Évaluer plusieurs modèles et comparer leur résultat à d'éventuels autres modèles connus.

LES DONNÉES

Tout part donc des données. Après avoir identifié le problème médical qui doit être résolu ou amélioré par une IA, il s'agit d'identifier les bases de données qui pourront être utilisées pour entraîner l'algorithme. La récolte et l'accès à des données médicales sont évidemment particulièrement sensibles. La législation européenne, appelée *Règlement général de protection des données* (RGPD), impose des normes très strictes de respect des données personnelles, ce qui inclut les données médicales. Ces normes sont un peu plus souples aux États-Unis et surtout en Chine. S'il est évidemment souhaitable de protéger de façon adéquate nos données, cette rigidité légale peut également être une entrave au développement d'analyses statistiques et à la création de modèles d'intelligence artificielle en Europe. Cet élément est important à prendre en compte dans toute démarche de création d'une intelligence artificielle dans la mesure où l'obtention de l'accès aux données sera très longue (de six mois à un an, ou plus selon les cas), à moins d'utiliser une base de données en accès libre et anonymisée sur Internet.

Une fois les données obtenues, il faut réaliser un travail bien souvent titanesque de nettoyage et de sélection des données pertinentes. Cette étape est primordiale : la qualité des données est très importante. On dit en anglais *garbage in, garbage out* : en utilisant des données de mauvaise qualité, le modèle d'IA sera lui aussi nécessairement mauvais. Ce « peaufinage » est le plus souvent réalisé dans le cadre d'une collaboration entre un expert du domaine (médecin) et un ingénieur informatique ou un chercheur. En effet, si l'ingénieur ou le chercheur a les capacités pour programmer ou coder, il n'a pas toujours les connaissances nécessaires afin de sélectionner les données pertinentes ou de repérer les éventuelles erreurs ou anomalies dans la base de données. Cette étape représente entre 70 et 80 % du temps nécessaire à la création d'une intelligence artificielle médicale.

Les chercheurs en neurosciences estiment que le cerveau humain est capable de prendre une décision adéquate en utilisant au maximum cinq facteurs. Au-delà, notre cerveau n'est plus capable de pondérer correctement les choses. Notre cerveau attribuera trop d'importance à un facteur anecdotique, ou pas assez à un facteur très important. Il ne sera pas capable de réaliser une prédiction ou d'avoir une intuition correcte. Or, dans le domaine médical, un médecin est bien souvent amené à prendre des décisions de diagnostic ou de thérapeutique sur un nombre de facteurs bien plus grand. C'est pour cette raison que l'IA peut avoir un intérêt.

Les différents types de données que l'on peut être amené à utiliser sont :

— Cliniques (état général du patient, symptômes, examen clinique, résultats de prises de sang).
— Biologie moléculaire (génomique, protéomique).
— Imagerie (taille d'une lésion, nombre, données métaboliques, radiomique).
— Thérapeutiques (dose de médicament, nombre de prises, intervention chirurgicale).

LA MÉTHODE D'ANALYSE

Vient ensuite le moment de choisir le type d'algorithme d'intelligence artificielle que l'on souhaite utiliser. Car il faut savoir que, même si on parle beaucoup du *Deep Learning*, il existe en réalité un très grand nombre de méthodes disponibles :

— Arbres de décisions (*Decision Tree*), où un algorithme simple crée des classes mutuellement exclusives en répondant à des questions prédéfinies.
— Classification naïve bayésienne (*Naive Bayes*), qui permet de classer des variables selon un modèle probabiliste.
— K plus proches voisins (*k-Nearest Neighbor*), où une entité est classifiée selon les caractéristiques de son plus proche voisin.
— Machines à vecteur de support (*Support Vector Machine*), où un algorithme classifiera des données en fonction d'un seuil linéaire.
— Réseaux neuronaux (*Artificial Neural Network*), qui utilisent des modèles inspirés des réseaux biologiques.
— Apprentissage profond (*Deep Learning*), une évolution des réseaux neuronaux qui comportent de multiples couches et sous-couches de neurones.

Très concrètement, on va fournir à l'un de ces algorithmes la base de données préparée dans la première étape. On lui spécifie à partir de quel facteur il doit travailler et quelle prédiction il doit réaliser : imaginons que je souhaite prédire le taux de guérison d'un cancer cinq ans après son diagnostic. La base de données comprend pour chaque patient un ensemble de facteurs médicaux et dans une autre colonne l'information de sa guérison ou non à cinq ans. Nous allons donner à l'algorithme l'ensemble des facteurs médicaux et lui demander de prédire la guérison. L'algorithme va alors apprendre à reconnaître les facteurs qui permettent de prédire la guérison. Il modifiera progressivement et automatiquement ses paramètres, afin de minimiser l'erreur entre sa prédiction et la réalité, jusqu'à obtenir le meilleur taux de réponses correctes possible. On dit que l'algorithme converge vers une solution optimale. Les méthodes les plus fréquemment utilisées (parce qu'elles donnent les meilleurs résultats) sont le *Deep Learning* (pour les images) et les *Decision Tree*, le plus souvent sous une forme *Gradient-Boosted* (*XGBoost*) pour les données tabulaires. Nous détaillons donc le fonctionnement de ces deux méthodes dans le chapitre suivant.

APPRENTISSAGE SUPERVISÉ OU NON SUPERVISÉ

Le but de l'apprentissage supervisé est de prédire une variable de sortie connue. Il doit d'abord être entraîné sur un ensemble de données où les caractéristiques initiales et finales sont labellisées et connues afin de créer un modèle qui classe le plus fidèlement possible ces données d'entraînement. Le modèle sera ensuite généralisé pour prédire les données finales (comme la survie globale, la guérison ou la toxicité) à partir des données d'entrée de n'importe quel patient. Dans l'apprentissage non supervisé, l'ensemble de données explorées n'est pas labellisé et c'est l'algorithme qui apprend et découvre des motifs de corrélation, sans entraînement initial et sans intervention humaine. Cette méthode permet de détecter des facteurs pronostiques ou prédictifs qui n'auraient pas pu être déterminés par un raisonnement humain.

VALIDATION DE L'IA SUR D'AUTRES DONNÉES

L'avant-dernière étape consiste à tester l'intelligence artificielle sur une nouvelle base de données, différente de celle qui était utilisée pour l'entraîner. On va donc demander à l'algorithme de réaliser une prédiction à partir de facteurs initiaux et l'on pourra comparer la prédiction réalisée à la réalité. Si les prédictions sont correctes, alors on dit que l'intelligence artificielle est généralisable. On pourra alors s'en servir éventuellement sur de nouveaux patients.

COMPARAISON AUX MÉDECINS OU À D'AUTRES IA

La dernière étape va permettre de comparer les performances de cette nouvelle IA à celles d'un médecin ou à celles d'autres IA déjà existantes. Pour cela, on prendra une base de données complète et on

évaluera les prédictions réalisées par la nouvelle IA, un médecin ou un groupe de médecins, ou une IA déjà existante. Il n'est cependant pas toujours possible de comparer les performances d'un médecin à celles d'une IA – typiquement pour les IA qui réalisent des tâches dont les humains ne sont pas capables (par exemple, prédire avec exactitude la guérison d'un cancer).

À l'inverse, cela peut tomber sous le sens, mais si le médecin, ou une IA qui existe déjà, produit de meilleurs résultats que l'IA que l'on vient de créer, alors on peut malheureusement, et tout simplement, mettre à la poubelle l'ensemble du travail réalisé !

Comment fonctionne un réseau neuronal ?

Dans un réseau neuronal, la reconnaissance d'un objet figurant sur une image ne repose pas sur un ensemble de règles. Ainsi, il n'est pas nécessaire d'avoir préétabli qu'un chien a quatre pattes et deux oreilles pour reconnaître un chien sur une image. Il est par contre nécessaire de fournir plusieurs dizaines de milliers d'images de chien au réseau neuronal qui apprendra de lui-même à différencier un chien d'un chat.

De façon simplifiée, un réseau neuronal est constitué d'une succession de couches de neurones. Chacun de ses neurones prend des valeurs d'entrée, y applique un poids, en fait la somme qui, si elle dépasse un certain seuil, déclenche l'activation du neurone et la transmission de l'information.

Chaque neurone et chaque couche sont reliés par des synapses. Un neurone reçoit ainsi les données issues de plusieurs neurones de la couche précédente et passera l'information à la couche suivante. Chaque neurone reçoit un signal et, si ce signal dépasse un certain seuil, il enclenche une fonction d'activation et passe le signal à la couche suivante. Chaque neurone possède un poids préétabli au hasard à l'initialisation du réseau, qui traduit son importance. Le but du réseau neuronal est d'ajuster les poids et les paramètres de ses neurones afin de minimiser l'erreur existant entre la prédiction ou la classification qu'il a réalisée et la réalité. L'entraînement consiste à passer un certain nombre de données (*batch*) dans le réseau neuronal plusieurs fois (*epoch*) afin d'en ajuster les paramètres pour obtenir le meilleur résultat possible.

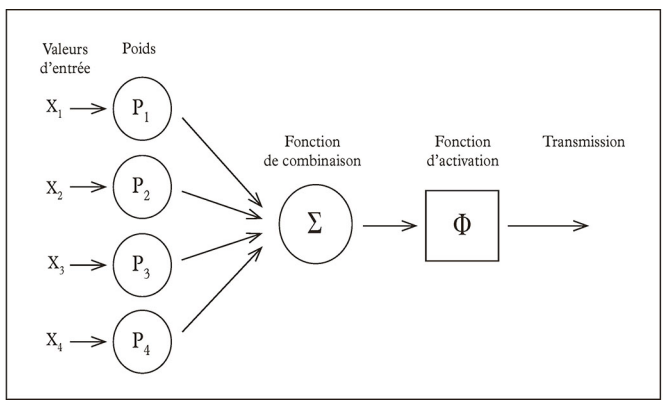


Schéma de fonctionnement d'un neurone.

Comment fonctionne un arbre décisionnel ?

Une des autres méthodes d'intelligence artificielle les plus populaires est l'arbre décisionnel et ses variantes de forêts aléatoires. Cette méthodologie, différente du *Deep Learning*, utilise un arbre de décision dans lequel les feuilles représentent les valeurs de la variable et où les nœuds sont les combinaisons de variables d'entrée qui mènent à ses valeurs.

À chaque embranchement, l'algorithme va apporter une réponse binaire et c'est l'ensemble des réponses de chaque embranchement qui fournira la prédiction ou la classification. Les règles qui régissent chacun de ces embranchements ne sont pas prédéfinies et c'est l'algorithme qui les crée lui-même lors de la phase d'apprentissage. Ce type d'algorithme est particulièrement adapté à l'analyse de données tabulaires. Il est par ailleurs beaucoup plus facilement compréhensible qu'un algorithme de *Deep Learning*, puisqu'il correspond à un ensemble de règles facilement interprétables. Il est par conséquent l'un des algorithmes les plus utilisés dans le cadre d'intelligence artificielle créée pour la médecine prédictive à partir de données écrites de dossiers médicaux, comme nous le verrons par la suite.

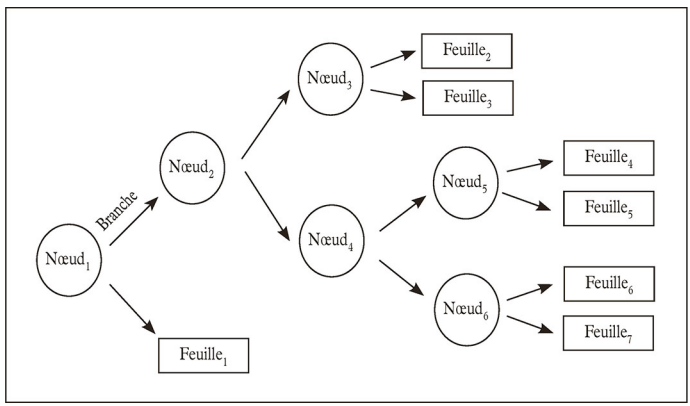


Schéma d'un arbre décisionnel.

Bases techniques du *Deep Learning*

Il est impossible de décrire ici de façon exhaustive les aspects techniques et les différentes architectures des réseaux neuronaux profonds. Cependant, vous trouverez dans cette partie certaines bases qui vous permettront d'aller plus loin et qui vous aideront à comprendre le fonctionnement concret d'un réseau neuronal. Cette partie détaille donc certains éléments clés des réseaux neuronaux et comment ils peuvent être combinés.

L'INITIALISATION

Lorsqu'un réseau de neurones artificiels est entraîné, on cherche à définir le poids de chaque neurone du réseau. La valeur initiale attribuée au poids de chaque neurone avant l'entraînement est donc très importante. On attribue en général une valeur aléatoire à chaque neurone, comprise entre – 1 et 1. Le réseau neuronal modifiera ensuite automatiquement le poids de chaque neurone à chaque cycle d'apprentissage en comparant la prédiction obtenue par une certaine combinaison de poids à la réalité des données fournies.

À l'inverse, lorsqu'une méthode de *transfer learning* est utilisée, les poids au début de l'apprentissage ont déjà été pré-entraînés sur une base de données similaire. Ainsi, un réseau neuronal entraîné à reconnaître un certain type d'objets sur des photos peut être réutilisé et partiellement réentraîné pour reconnaître d'autres objets sur d'autres photos. Les poids pré-entraînés seront alors un meilleur point de départ pour l'apprentissage. Cette méthode est particulièrement adaptée dans le cas où l'on dispose de peu de données d'apprentissage.

LES FONCTIONS D'ACTIVATION – *Activation Function*

Concrètement, un neurone artificiel reçoit une information en amont, et la traite *via* une fonction d'activation, ce qui permettra d'activer le neurone et de passer l'information, ou de le désactiver et de ne pas passer l'information.

Les fonctions d'activation sont inspirées des potentiels d'action que l'on voit dans un véritable neurone. Un potentiel d'action est un signal électrique qui parcourt un neurone et qui va entraîner la libération ou la rétention de neurotransmetteurs dans les synapses, c'est-à-dire dans les connexions existantes avec d'autres neurones. Alors que les potentiels d'action dans les cellules sont binaires (la cellule est soit activée, soit désactivée), les neurones artificiels peuvent fournir un signal de sortie binaire ou une valeur numérique. Ils sont également plus flexibles sur la méthode de sortie du signal et permettent par exemple de calculer la somme d'une sortie, sa dérivée (taux de changement) ou des fonctions plus complexes. De plus, un neurone biologique ne peut être activé qu'à une certaine fréquence prédéfinie car il possède une période réfractaire durant laquelle il ne peut plus être activé quel que soit le signal d'entrée. Les neurones artificiels n'ont évidemment pas cette limitation.

Il existe un grand nombre de fonctions d'activation différentes qui partagent plusieurs propriétés : elles sont non linéaires, différentiables et leur signal de sortie a un intervalle prédéfini. Parmi les fonctions d'activation les plus usuelles, on retrouve les fonctions binaire, linéaire, sigmoïde et l'unité de rectification linéaire (*ReLu*). Chaque neurone d'un réseau n'a pas obligatoirement la même fonction d'activation. La plupart du temps, une fonction d'activation est choisie pour toutes les couches profondes et c'est une fonction plus simple, par exemple binaire, qui est programmée pour les neurones de la couche de sortie.

L'EXTINCTION DE NEURONES – *Dropout*

L'extinction de neurones est une technique de régularisation qui permet d'éviter le surapprentissage (*overfitting*), phénomène qui se produit lorsqu'un réseau neuronal apprend les variables spécifiques d'une base de données, qui ne sont pas retrouvées dans une autre base. Le réseau sera alors incapable d'accomplir sa tâche sur une base de données différente de celle utilisée pour son entraînement et sera donc totalement inutile.

L'extinction de neurones consiste à désactiver de façon aléatoire à chaque cycle d'apprentissage un certain pourcentage des neurones d'une couche. On utilise en général un ratio de 50 % d'extinction à chaque cycle d'apprentissage.

DIFFÉRENTS TYPES DE COUCHES

Couche entièrement connectée – Fully Connected Layer

On parle d'une couche entièrement connectée lorsque tous les neurones d'une couche se connectent à tous les neurones de la couche suivante. On parle aussi de couche *dense*. Chaque couche est constituée d'un nombre arbitraire de neurones qui reçoivent un signal d'entrée pondéré de la couche précédente, la traite avec la fonction d'activation et passent en sortie le fruit de ce traitement mathématique.

Couche convolutive – Convolutional Layer

La convolution consiste à définir un noyau convolutif (c'est-à-dire une matrice dont la dimension sera adaptée aux données d'entrée), qui sera déplacé sur l'ensemble de l'image afin de l'analyser. Pour analyser une image par exemple de 256 × 256 pixels, on peut définir une matrice de 2 × 2 pixels qui analysera donc les pixels de l'image quatre par quatre. Ce type d'approche est particulièrement utilisé dans la reconnaissance d'objets naturels.

Couche de mise en commun – Pooling Layer

Ce type de couche permet de combiner les résultats de différentes couches en amont et de différents neurones en réalisant une somme, ou en gardant uniquement la valeur maximale ou minimale, ou bien encore en en faisant la moyenne. Cela permet notamment de simplifier la série de valeurs numériques extraites d'une image pour l'analyse.

*Couche d'*Inception

L'*Inception* consiste à utiliser différents filtres de différentes tailles sur l'image d'entrée et ensuite d'ajouter les résultats obtenus. Ce processus permet d'obtenir des informations à différentes échelles et donc d'extraire des variables à différents niveaux de façon simultanée. Typiquement, cela permet sur une image d'analyser les motifs mais aussi les couleurs ou les bords des objets qu'elle contient.

Bloc résiduel – Residual Block

Un bloc résiduel est constitué de deux ou trois couches entièrement connectées de neurones qui ont une connexion de « court-circuitage » entre la couche d'entrée et la couche de sortie. Ce court-circuit permet au réseau de réduire le nombre de couches à utiliser et donc de réduire le risque de *surapprentissage*. En effet, si une couche de neurones ne fournit pas de meilleurs résultats dans la prédiction que la couche qui la précède, elle est alors court-circuitée et le signal est directement passé à la couche de neurones suivante.

DIFFÉRENTS TYPES DE RÉSEAUX

Réseau entièrement connecté – Fully Connected Network

Ce type de réseau correspond à plusieurs couches entièrement connectées utilisant bien souvent la fonction d'activation *ReLu* et une extinction des neurones (*dropout*). C'est le type de réseau le plus simple mais il n'est pas toujours approprié pour analyser une image.

Réseau convolutif – Convolutional Neural Network

Les réseaux convolutifs sont une des formes les plus utilisées de réseaux neuronaux profonds. Ils sont particulièrement adaptés à l'analyse d'image. Ils sont constitués d'une association de couches entièrement connectées et de couches convolutives. Ils permettent notamment d'« écraser » une image 2D en une série de valeurs unidimensionnelles, plus facile à analyser. Concrètement, les couches convolutives apprennent les caractéristiques de l'image qui sont importantes et les couches entièrement connectées apprennent quant à elles les poids et les combinaisons de ces caractéristiques les plus importantes pour réaliser une prédiction correcte.

Réseau en U et réseau en V – U-Net *et* V-Net

Ces réseaux tiennent leur nom de la forme globale de leur architecture. Le rôle des premières couches (qui descendent vers le bas du U) est de réduire la résolution d'une image et d'en extraire les caractéristiques clés. La deuxième partie du réseau (qui remonte vers la partie droite du U) permet de reconstituer l'image dans sa résolution native tout en en gardant les caractéristiques d'intérêt. Ce type de réseau est principalement utilisé pour segmenter automatiquement une image. La segmentation est l'une des étapes nécessaires pour déterminer la place d'un objet dans une image. Ces réseaux sont donc particulièrement intéressants pour l'analyse automatique d'imagerie médicale.

Réseau convolutif densément connecté – DenseNet

Dans ce type de réseau, les couches convolutives sont toutes connectées les unes aux autres. Ainsi, sur un réseau à cinq couches, la couche numéro 1 sera connectée directement aux 2e, 3e, 4e et 5e, et ainsi de suite pour les autres couches. Ce type de schéma a été développé car certains réseaux étaient si profonds que l'information extraite dans les premières couches était perdue. Un *DenseNet* permet au contraire à l'information de s'écouler (*flow*) à travers l'ensemble du réseau, aussi profond soit-il.

Réseau antagoniste génératif – Generative Adversarial Network

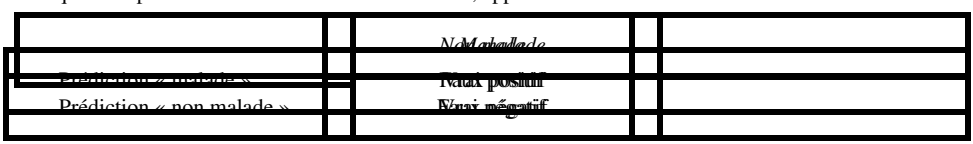
Le principe du réseau antagoniste génératif a été décrit en 1990 par Jürgen Schmidhuber sous le nom d'*Artificial Curiosity*, puis popularisé en 2014 par Ian Goodfellow, également chercheur en *Machine Learning*, devenu directeur de la recherche en IA d'Apple depuis, avant d'en démissionner en 2022. Dans ce modèle, deux réseaux sont en compétition : le premier réseau, appelé générateur, doit créer une information (souvent une image), tandis que le deuxième réseau, son adversaire appelé discriminateur, essaye de déterminer si l'information générée est réelle ou si elle est le résultat du générateur. Ce type de réseau est principalement utilisé pour générer des images artificielles, par exemple des photos de visage. En médecine, il pourrait être détourné pour générer les images virtuelles de scanner ou IRM, pour réaliser par exemple ce que l'on appelle des attaques antagoniste.

À quoi reconnaît-on une bonne intelligence artificielle ?

Les résultats obtenus par une IA doivent être interprétés avec prudence. Il est ainsi nécessaire de connaître certains critères de qualité afin de pouvoir juger de la qualité réelle d'un algorithme. Dans cette partie, je vous propose de définir les principaux critères actuellement utilisés. Ce chapitre sera particulièrement utile pour les lecteurs voulant aller plus loin en lisant les études qui apparaissent dans les références de bas de page. Cependant, afin de faciliter et de fluidifier la lecture, je ne reprendrai pas systématiquement pour chaque étude les performances chiffrées des algorithmes cités.

ALGORITHMES DE CLASSIFICATION

Ce type d'algorithme est utilisé afin de ranger ou classifier une image, un symptôme, des résultats biologiques, ou plus globalement un patient. Comme pour tout test diagnostique, quelques critères de qualité et leur définition doivent être compris. Prenons l'exemple d'un algorithme qui permettrait de classer un patient en malade ou non malade. Il réalise une prédiction que l'on peut comparer à la vérité, ce qui nous permet d'obtenir le tableau ci-dessous, appelé matrice de confusion.



On obtient donc :
VP (vrais positifs) : le nombre d'individus malades avec une prédiction malade.
FP (faux positifs) : le nombre d'individus non malades avec une prédiction malade.
FN (faux négatifs) : le nombre d'individus malades avec une prédiction non malade.
VN (vrais négatifs) : le nombre d'individus non malades avec une prédiction non malade.

Sensibilité, ou recall
Probabilité que la prédiction soit « malade » si la maladie est bien présente. Elle s'obtient en calculant :

VP / (VP + FN).

Spécificité
Probabilité d'obtenir une prédiction négative chez les non malades. Elle se calcule avec la formule

VN / (VN + FP).

Valeur prédictive positive
Probabilité que le patient soit malade, comme prédit par le modèle.

Valeur prédictive négative
Probabilité que le patient ne soit pas malade, comme prédit par le modèle.

Classification Accuracy
Rapport entre le nombre de vrais positifs et le nombre total d'observations, soit :

(VP + VN) / (VP + FP + FN + VN).

Précision – Precision

Rapport des vrais positifs sur le nombre total de prédictions positives, soit :

$$VP / (VP + FP).$$

F1-score

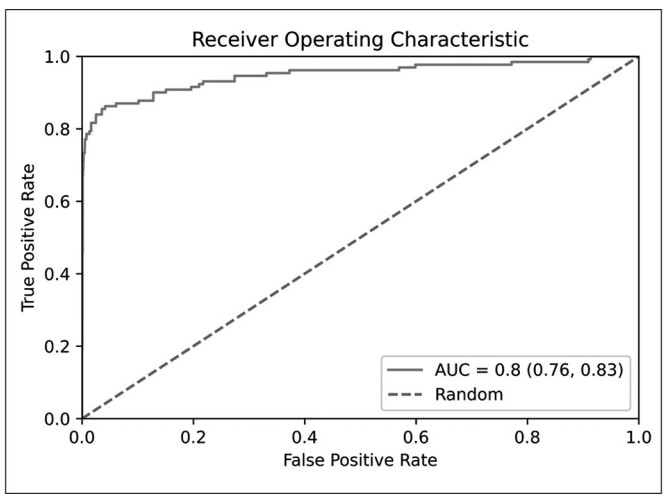
Moyenne harmonique de la *precision* et du *recall*, soit :

$$2 \times \text{precision} \times \text{recall} / (\text{precision} + \text{recall}).$$

Tous ces scores doivent tendre vers 1 (ou 100 %) pour que l'on puisse considérer l'algorithme comme bon.

Courbe ROC – Receiver Operating Characteristic curve

Pour comprendre une courbe ROC, il faut comprendre qu'un algorithme de classification calcule une probabilité d'appartenir au groupe malade ou non malade. En fonction du seuil de probabilité au-dessus duquel vous classez le patient comme malade (en général si la probabilité calculée est supérieure à 50 %), les résultats de votre algorithme seront très différents. La courbe ROC vous permet de voir les performances de votre algorithme en fonction de différents seuils de positivité.



Exemple d'une courbe ROC. La ligne en pointillé montre un algorithme qui serait aussi mauvais qu'une pièce lancée à pile ou face pour prédire la maladie. La courbe montre les performances d'une IA. Idéalement, la courbe doit aller le plus haut possible dans le coin supérieur gauche.

Aire sous la courbe – AUC (Area Under the Curve)

Mesure de l'aire sous la courbe ROC.

Courbe precision-recall – Precision-recall curve

Comme pour la courbe ROC, la courbe *precision-recall* montre ces deux critères en fonction de différents seuils de positivité. Elle est particulièrement utile lors de problèmes de déséquilibre des classes de patients, lorsque vous avez peu de patients malades dans une base de données d'entraînement, ce qui peut fausser les résultats et leur interprétation.

Aire sous la courbe precision-recall – Precision-recall Area under the curve

Mesure de l'aire sous la courbe *precision-recall.*

ALGORITHMES DE RÉGRESSION

Les algorithmes de régression permettent de prédire une valeur continue et non pas une catégorie. Ils sont par exemple utilisés pour prédire la durée de survie ou une fréquence.

Erreur quadratique moyenne – Mean Squared Error
: Mesure de la différence au carré entre la valeur des prédictions et la réalité.

Erreur absolue moyenne – Mean Absolute Error
: Mesure de la différence absolue entre la valeur des prédictions et la réalité.

Index des concepts

Anatomopathologie : 91, 93, 100, 138, 148, 172.
Arbre décisionnel : 183, 188, 189.
Bien-être : 161.
Cancérologie : 13, 14, 78, 79, 82.
Cardiologie : 96, 105, 146.
Chirurgie : 107, 108, 165, 169.
Dermatologie : 94, 95, 127.
Épidémiologie : 67, 74, 75, 79.
Gastro-entérologie : 98.
Médecine reproductive, gynécologie et obstétrique : 111.
Néphrologie : 113.
Neurologie : 105, 146.
Ophtalmologie : 54, 77.
Pédiatrie : 113.
Psychiatrie : 12, 13, 101, 103.
Radiologie : 83, 86, 87, 89, 92, 105, 138, 148.
Radiothérapie : 108.
Réseau neuronal : 32, 39, 40, 41, 46, 78, 94, 95, 99, 108, 111, 186, 187, 190, 191, 192.
— convolutif (CNN) : 78, 100, 195, 196.
— profond (*Deep Learning*) : 48, 49, 84, 112, 118, 131.
Urologie : 100.

Remerciements

Ingrid, Anaïs, Clémence, Noémie, Marie-Hélène, Henri, Pierre, Véronique, Bruno, Vincent, Caroline, Catherine, Martin, Émilie, toute l'équipe de l'HEGP, Lei, Maxime, Anita, Maximilien, Emmanuel, Alexandre, Thomas, Lucie, Pierre-François, David, Stanley.

[1]. Pour la définition du réseau neuronal, voir *infra*, page 186.

[2]. Pour de plus amples explications, voir l'annexe « Comment fonctionne un réseau neuronal », *infra*, page 186.

[3]. « Google's secret cache of medical data includes names and full details of millions – whistleblower », *The Guardian*, 12 novembre 2019.

[4]. Bibault J.-E., Bassenne M., Hongyi R., et al., *Deep Learning Prediction of Cancer Prevalence from Satellite Imagery*, Cancers, 2020.

[5]. V. Gulshan, L. Peng, M. Coram, M.C. Stumpe, D. Wu, et al*., Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs*, JAMA, 2016.

[6]. Bibault J.-E., Hancock S., Buyyounouski M., et al., *Development and Validation of an Interpretable Artificial Intelligence Model to Predict 10-Year Prostate Cancer Mortality*, Cancers, 2021.

[7]. https://prostatecancersurvival.herokuapp.com.

[8]. Sun R., Limkin E.J., Vakalopoulou M., et al., *A radiomics approach to assess tumour-infiltrating CD8 cells and response to anti-PD-1 or anti-PD-L1 immunotherapy : an imaging biomarker, retrospective multicohort study*, The Lancet Oncology, 2018.

[9]. Rajpurkar P, Irvin J., Zhu K. et al., *CheXNet : Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning,* arXiv, 2017.

[10]. Ahmed F., Bukhari S., Keshtkar F., *A Deep Learning Approach for COVID-19 8 Viral Pneumonia Screening with X-ray Images*, Digital Government : Research and Practice, 2021.

[11]. Ardila D., Kiraly A.P., Bharadwaj S., et al., *End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography*, Nature Medicine, 2019.

[12]. Oakden-Rayner L., Carneiro G., Bessen T. et al., *Precision Radiology: Predicting longevity using feature engineering and deep learning methods in a radiomics framework*, Scientific Reports, 2016.

[13]. Karimi D., Nir G., Fazli L., et al., *Deep Learning-based gleason grading of prostate cancer from histopathology images role of multiscale decision aggregation and data augmentation*, IEEE J Biomed Health Inform, 2020.

[14]. Esteva A., Kuprel B., Novoa R., et al., *Dermatologist-level classification of skin cancer with deep neural networks*, Nature, 2017.

[15]. Li J., Rapin J., Rosier A., et al., *Deep Neural networks improve atrial fibrillation detection in Holter : first results*, European Journal of Preventive Cardiology, 2016.

[16]. Aoki T., Yamada A., Aoyama K., et al., *Clinical usefulness of a deep learning-based system as the first screening on small-bowel capsule endoscopy reading*, Dig Endosc, 2020.

[17]. Weizenbaum J., *ELIZA – a computer program for the study of natural language communication between man and machine,* Commun ACM, 1966.

[18]. Bibault J.-E., Chaix B., Guillemassé A., et al., *A Chatbot Versus Physicians to Provide Information for Patients With Breast Cancer : Blind, Randomized Controlled Noninferiority Trial*, JMIR, 2019.

[19]. Eichstaedt J.C., Smith R.J., Merchant R.M., et al., *Facebook language predicts depression in medical records*, PNAS, 2018.

[20]. Reece A.G., Danforth C.M., *Instagram photos reveal predictive markers of depression*, EPJ Data Science, 2017.

[21]. Muehlematter U.J., Daniore P., Vokinger K.N., *Approval of artificial intelligence and machine learning based medical devices in the USA and Europe (2015-20) : a comparative analysis*, Lancet Digital Health, 2021.

[22]. Wu E., Wu K., Daneshjou R., et al., *How medical AI devices are evaluated : limitations and recommendations from an analysis of FDA approvals*, Nature Medicine, 2021.

[23]. Jin A., Yeung S., Jopling J., et al., *Tool Detection and Operative Skill Assessment in Surgical Videos Using Region-Based Convolutional Neural Networks*, arXiv, 2018.

[24]. Liang X., Bibault J.E., Leroy T., et al., *Automated contour propagation of the prostate from*

pCT to CBCT images via deep unsupervised learning, Medical Physics, 2021.

[25]. Pegah Khosravi, Ehsan Kazemi, Qiansheng Zhan, et al., *Deep Learning enables robust assessment and selection of human blastocysts after in vitro fertilization*, NPJ Digital Medicine, 2019.

[26]. Artzi N.S., Shilo S., Hadar E., et al., *Prediction of gestational diabetes based on nationwide electronic health records*, Nature Medicine, 2020.

[27]. Tomasev N., Glorot X., Rae J.W., et al., *A clinically applicable approach to continuous prediction of future acute kidney injury*, Nature, 2019.

[28]. Anselma L., Mazzei A., De Michieli F., *An artificial intelligence framework for compensating transgressions and its application to diet management*, Journal of Biomedical Informatics, 2017.

[29]. Zeevi D., Korem T., Zmora N., et al., *Personalized Nutrition by Prediction of Glycemic Responses*, Cell, 2015.

[30]. Hassoon A., Schrack J., Naiman D., et al., *Increasing physical activity amongst overweight and obese cancer survivors using an Alexa-based intelligent agent for patient coaching : protocol for the physical activity by technology help (PATH) trial*, JMIR Res Protoc, 2018.

[31]. Wen-Sheng Tseng V., Valliappan N., Ramachandran V., et al., *Digital biomarker of mental fatigue*, NPJ Digital Medicine, 2021.

[32]. Ajay V.S., Jindal D., Roy A., et al., *Development of a smartphone-enabled hypertension and diabetes mellitus management package to facilitate evidence-based care delivery in primary healthcare facilities in India : the mPower Heart Project*, J Am Heart Assoc, 2016.

[33]. Denis F., Lethrosne C., Pourel N., et al., *Randomized Trial Comparing a Web-Mediated Follow-up With Routine Surveillance in Lung Cancer Patients*, JNCI, 2017.

[34]. Yeung S., Rinaldo F., Jopling J., et al., *A computer vision system for deep learning-based detection of patient mobilization activities in the ICU*, NPJ Digital Medicine, 2019.

[35]. Singh A., et al., *Automatic detection of hand hygiene using computer vision technology*, J. Am Med Inform Assoc, 2020.

[36]. Chiu C.C., Tripathi A., Chou K., et al., *Speech recognition for medical conversations,* arXiv, 2018.

[37]. Chakraborty D., Yasir T., Zixu Y., *Assessment and prediction of negative symptoms of schizophrenia from RGB + D movement signals*, Proc. 19th International Workshop on Multimedia Signal Processing, 2017.

[38]. Obermeyer Z., Powers B., Vogeli C., et al., *Dissecting racial bias in an algorithm used to manage the health of populations*, Science, 2019.

[39]. Wang Y., Kosinski M., *Deep Neural Networks can detect sexual orientation from faces*, PsyarXiv, 2017.

[40]. Finalyson S.G., Bowers J.D., Ito J., et al., *Adversarial attacks on medical machine learning*, Science, 2019.

[41]. Roberts M., Driggs D., Thorpe M., et al., *Common pitfalls and recommendations for using machine learning to detect and prognosticate for COVID-19 using chest radiographs and CT scans*, Nature Machine Intelligence, 2021.

[42]. McDermott M.B.A., Wang S., Marinsek N., et al., *Reproducibility in machine learning for health research : Still a ways to go*, Science, 2021.

[43]. Copeland R., « Google's "Project Nightingale" gathers personal health data on millions of Americans », *Wall Street Journal*, 2019.

[44]. Khullar D., Casalino L.P., Qian Y., et al., *Public vs physician views of liability for artificial intelligence in health care*, JAMIA, 2021.

[45]. *Infra*, page 179.

[46]. Zhang D., Maslej N., Brynjolfsson E., et al., *The AI Index 2022 Annual Report, AI Index Steering Committee*, Stanford Institute for Human-Centered AI, Stanford University, mars 2022.

[47]. Zwetsloot R., Zhang B., Dreksler N., et al., *Skilled and Mobile : Survey Evidence of AI Researchers'Immigration Preferences*, arXiv, 2021.

ÉDITIONSDESÉQUATEURS

www.editionsdesequateurs.fr